



Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft Nr. 846

Johann Christian Rose

**Automatische, hochaufgelöste 3D-Phänotypisierung von
Trauben und Beeren der Weinrebe unter Feldbedingungen**

München 2020

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325

ISBN 978-3-7696-5258-1

Diese Arbeit ist gleichzeitig veröffentlicht in:

**Schriftenreihe des Instituts für Geodäsie und Geoinformation der Rheinischen Friedrich-Wilhelms
Universität Bonn, ISSN 2699-6685, Nr. 61, Bonn 2020**



Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft Nr. 846

Automatische, hochaufgelöste 3D-Phänotypisierung von Trauben und Beeren der Weinrebe unter Feldbedingungen

Inaugural-Dissertation zur

Erlangung des Grades

Doktor-Ingenieur (Dr.-Ing.)

der Hohen Landwirtschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms Universität Bonn

vorgelegt im August 2019 von

Johann Christian Rose, M.Sc.

Aus Olpe

München 2020

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325

ISBN 978-3-7696-5258-1

Diese Arbeit ist gleichzeitig veröffentlicht in:

Schriftenreihe des Instituts für Geodäsie und Geoinformation
der Rheinischen Friedrich-Wilhelms Universität Bonn,
ISSN 2699-6685, Nr. 61, Bonn 2020

Adresse der DGK:



Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 • D – 80 539 München
Telefon +49 – 331 – 288 1685 • Telefax +49 – 331 – 288 1759
E-Mail post@dgk.badw.de • <http://www.dgk.badw.de>

Prüfungskommission:

Referent: Prof. Dr.-Ing. Heiner Kuhlmann

Korreferent: Jun.-Prof. Dr. Ribana Roscher

Korreferent: Prof. Dr. Reinhard Töpfer

Tag der mündlichen Prüfung: 31.10.2019

© 2020 Bayerische Akademie der Wissenschaften, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen.

Automatische, hochaufgelöste 3D-Phänotypisierung von Weintrauben und Beeren der Weinrebe unter Feldbedingungen

Zusammenfassung

Pilzwiderstandsfähige Sorten (PIWI) sind eine der Innovationen im Weinbau. Das Projekt *Novel Viticulture Systems for sustainable production and products (NoViSys)* kombiniert die PIWIs zur Effizienzsteigerung mit einem neuen Erziehungssystem für Weinreben, dem Minimalschnitt im Spalier (MSS) und vergleicht es mit der Erziehung im Traditionellen Spalier System (TS). Ein wichtiges phänotypisches Merkmal zum Vergleich beider Erziehungssysteme ist der Ertrag. Für vergleichende Untersuchungen und Optimierungsmaßnahmen soll der Ertrag während der Saison in mehreren Reifestadien BBCH75 und BBCH89 ermittelt werden. Die Trauben müssen daher intakt an den Weinreben verbleiben. Um den Ertrag nicht-invasiv zu bestimmen, werden die sog. *Ertragsparameter* ermittelt: die Anzahl der Trauben und Beeren, sowie die Beerengröße.

Traditionellerweise werden die Ertragsparameter durch einen erfahrenen Züchter aus wenigen visuell erhobenen Stichproben extrapoliert und zur Ertragsschätzung genutzt. Sensorbasierte Methoden zielen darauf ab, im Hochdurchsatz auf automatische und nicht-invasive Weise, objektive und genaue Ertragsparameter von einer Vielzahl von Weinreben zu bestimmen. Ziel ist es, durch eine größere und vollständigere Stichprobe der Ertragsparameter die statistische Sicherheit der Ertragsschätzung zu erhöhen. Die Aufgabe dieser Dissertation lautet daher:

Entwicklung einer nicht-invasiven, sensorgestützten Hochdurchsatz-Methode zur automatischen Schätzung der Ertragsparameter und eine auf den Ertragsparametern basierende Schätzung des Ertrages für TS- und MSS-Erziehungssysteme im Reifestadium BBCH75 und BBCH89 unter Feldbedingungen.

In dieser Arbeit wird eine auf RGB-Kameras basierende Lösung der Aufgabe vorgestellt und evaluiert. Drei **Teilaufgaben zur Schätzung der Ertragsparameter und des Ertrages** werden dazu gelöst:

1) Datenerhebung: Entwicklung und Untersuchung mobiler Sensorplattformen für die Datenerhebung im Hochdurchsatz unter Feldbedingungen für beide Erziehungssysteme.

Zwei mobile Sensorplattformen werden untersucht. Auf ihnen montierte RGB-Kameras nehmen in Bewegung automatisch Bilder der Weinreben aus multiplen Perspektiven entlang der Horizontalen und der Vertikalen auf. Aus den über GPS georeferenzierten RGB-Bildern werden mittels einer Multi-View-Stereo Software farbige, metrisch skalierte Punktwolken der Weinreben mit hoher geometrischer Auflösung rekonstruiert.

2) Datenklassifikation: Entwicklung einer geeigneten algorithmenbasierten Klassifikationsstrategie, um in den Daten Beeren und Trauben von Blättern, Ästen und anderen Bestandteilen automatisch zu unterscheiden.

Zur Klassifikation wird der überwachte Klassifikator Import Vector Machine genutzt. Mehrere geometrie- und farbbasierte Deskriptoren werden hinsichtlich der Klassifikationsgenauigkeit untersucht und der beste Deskriptor identifiziert. Räumliche Beziehungen zwischen den 3D-Punkten der klassifizierten Punktwolken werden in einem Graph-Cut Ansatz zur Verbesserung der Klassifikationsergebnisse ausgenutzt.

3) Quantifikation: Entwicklung einer algorithmenbasierten Strategie zur automatischen Schätzung der Ertragsparameter und des Ertrages aus den klassifizierten Daten.

Die Untersuchungen umfassen die Schätzung der Ertragsparameter für beide Erziehungssysteme und Reifestadien, sowie einen Vergleich mit Referenzdaten. Die Ertragsparameter werden anhand der klassifizierten 3D-Punkte durch ein 3D-Connected-Components Verfahren in Kombination mit Kugelapproximationen, sowie Kalibrierwerten zur Korrektur nicht detektierter und unsichtbarer Beeren und Trauben bestimmt. Des Weiteren wird eine Funktionsvorschrift zur Schätzung des Ertrages auf Basis der geschätzten Ertragsparametern formuliert. Die Ertragsschätzung wird für beide Erziehungssysteme im BBCH89 Reifestadium vorgenommen. Zwischen geschätztem und tatsächlichem Ertrag werden Abweichungen von 5% für das TS-Erziehungssystem und -21% für das MSS-Erziehungssystem erreicht.

Automatic, high-resolution 3D-phenotyping of grapevine bunches and berries under field conditions

Abstract

Fungus resistant grape varieties (PIWI) are one of the innovations in viticulture. The project *Novel Viticulture Systems for sustainable production and products (NoViSys)* combines the PIWIs with the new semi-minimal pruned hedge (SMPH) training system to further increase their efficiency and compares their training in the SMPH system to the traditional vertical shoot-positioned (VSP) training system. One important phenotypic trait to compare both training systems by is the yield. For comparative studies and optimization procedures the yield has to be determined multiple times during the season for the ripeness stages BBCH75 and BBCH89. For this reason the grape bunches have to remain intact and attached to the grape vines. To predict the yield non-invasively the so called *yield-parameters* are determined: the number of grape bunches, berries and the size of the berries.

Traditionally the yield parameters are extrapolated by an experienced breeder from a few visually determined samples which are then used to predict the yield. Sensor based methods aim at predicting objective and accurate yield parameters from a multitude of grape vines automatically and non-invasively. The purpose is to increase the statistic reliability of the prediction through a larger and more complete sample of the yield parameters. The objective of this dissertation is thus as follows:

Development of a high-throughput, non-invasive, sensor-based method for the automated prediction of yield parameters and a yield prediction based on the yield parameters for the VSP and SMPH training system in the ripeness stages BBCH75 and BBCH89 under field conditions.

In this work a solution based on the usage of RGB-Cameras is presented and evaluated. Three *subtasks to predict the yield parameters and the yield* are solved:

1) Data Acquisition: Development and Evaluation of mobile sensor platforms for high-throughput data acquisition under field conditions for both training systems.

The performance of two mobile sensor platforms is studied. While moving RGB-cameras mounted on the platforms take images from multiple vertical and horizontal perspectives toward the grape vine rows. From the images georeferenced by GPS colored metric point clouds with a high geometric resolution are reconstructed using a Multi-View-Stereo software.

2) Data classification: Development of an algorithm driven *classification strategy* to automatically differentiate berries and grape bunches from leaves, branches and other objects present in data.

Classification is done using the Import Vector Machine classifier. Multiple geometry- and color-based feature descriptors are studied regarding classification accuracy and the best feature descriptor is identified. Spatial relationships between the 3D-points of the classified point clouds are exploited in a graph-cut approach to improve the classification results.

3) Quantification: Generation and evaluation of an algorithm driven method to *automatically determine the yield parameters and predict the yield*. from the classified data.

The experiments encompass the determination of yield parameters for both training systems and for both ripeness stages as well as a comparison with reference data. Yield parameters are determined from the classified 3D-points through a 3D connected-components method combined with sphere approximation and calibration parameters to correct for non-detected and invisible berries and grape bunches. Further a function to predict the yield based on the yield parameters is defined. Yield prediction is done for both training systems in the BBCH89 ripeness stage. Deviations of 5% for the VSP and -21% for the SMPH training system are achieved.

Inhaltsverzeichnis

Vorwort	1
1 Einleitung	3
1.1 Projekt NoViSys	3
1.2 Aufgabe	8
1.3 Aufbau der Arbeit	9
2 Stand der Forschung	11
2.1 Bildbasierte 2D-Methoden	11
2.2 3D-Methoden	14
2.3 Beitrag dieser Arbeit	15
2.3.1 Einordnung der Fachbeiträge	15
2.3.2 Vorgehensweise zur Lösung der drei Teilaufgaben	19
3 Methodische Grundlagen	21
3.1 3D-Rekonstruktion der Anbaureihen mittels einer SfM/MVS-Methode	21
3.1.1 Bildpunkte und innere Kameraparameter	21
3.1.2 Structure-from-Motion anhand homologer Bildpunkte	22
3.1.3 Multi-View-Stereo: Verdichtung der dünnen Punktwolke	24
3.1.4 Bewertung der Punktwolke anhand des Rückprojektionsfehlers	25
3.1.5 Zusammenfassung	26
3.2 Datenklassifikation	26
3.2.1 Grundlegende Klassifikationsstrategie	26
3.2.2 Erläuterung der untersuchten Deskriptoren	28
3.2.3 Überwachte vs unüberwachte Klassifikation	36
3.2.4 Erstellung von Trainings- und Testdaten	36
3.2.5 Überwachte Klassifikatoren	37
3.2.6 Graph-Cut Software GCO	40
3.2.7 Zusammenfassung	43

4 Praxisbeispiel 1: Voruntersuchungen mit dem PHENObot	45
4.1 Datenerhebung mit dem PHENObot	45
4.1.1 PHENObot	45
4.1.2 Bildaufnahme mit dem PHENObot	46
4.1.3 3D-Rekonstruktion der Anbaureihen	49
4.1.4 Vorverarbeitung der Punktwolken	51
4.1.5 Detailgrad der Punktwolken	54
4.1.6 Zusammenfassung	55
4.2 Datenklassifikation	56
4.2.1 Evaluierung der Deskriptoren	56
4.2.2 Weiterführende Untersuchungen	62
4.2.3 Klassifikation von 5m Ausschnitten beider Erziehungssysteme und Reifestadien . . .	66
4.3 Quantifikation	71
4.3.1 Detektion sichtbarer Beeren und Trauben	71
4.3.2 Evaluation der Auswertemethode	83
4.4 Zwischenfazit: Voruntersuchungen mit dem PHENObot	96
5 Praxisbeispiel 2: Ertragsschätzung mit dem Phenoliner	97
5.1 Datenerhebung	97
5.1.1 Der Phenoliner und seine Komponenten	97
5.1.2 Bildaufnahme mit dem Phenoliner	103
5.1.3 3D-Rekonstruktion der Anbaureihen	106
5.1.4 Vorverarbeitung der Punktwolken	108
5.1.5 Detailgrad der Punktwolken	109
5.1.6 Zusammenfassung	111
5.2 Datenklassifikation	112
5.2.1 Ermittlung optimaler Radien für den SFHc-Deskriptor	112
5.2.2 Klassifikation der fünf Abschnitte beider Erziehungssysteme	113
5.3 Quantifikation: Ertragsschätzung aus Ertragsparametern	114
5.3.1 Ablauf der Schätzung der Ertragsparameter und des Ertrages	116
5.3.2 Berechnung der Kalibrierwerte	118
5.3.3 Schätzung der Gesamtanzahl aller Beeren und Trauben	128
5.3.4 Funktionsvorschrift zur Schätzung des Ertrages	129
5.3.5 Schätzung des Ertrages der fünf Abschnitte beider Erziehungssysteme	129
5.3.6 Diskussion: Ertragsschätzung im TS- und MSS-Erziehungssystem im BBCH89 Reifestadium	132

6	Fazit und Ausblick	137
6.1	Fazit	137
6.2	Ausblick	139
	Abkürzungsverzeichnis	145
	Literaturverzeichnis	147

Vorwort

Im Zuge der vorliegenden Dissertation wurden folgende Veröffentlichungen verfasst, an denen der Autor dieser Arbeit als Erstautor den wesentlichen Beitrag geleistet hat, mit Ausnahme der letzten Veröffentlichung bei der er als Koautor mitwirkte:

- Rose, J.C.; Paulus, S.; Kuhlmann, H. Accuracy Analysis of a Multi-View Stereo Approach for Phenotyping of Tomato Plants at the Organ Level. *Sensors*, 15, 9651-9665, 2015. (peer-reviewed)
- Rose, J. C. & Kuhlmann, H. Classification of plant organs in 3D point clouds based on geometry and color descriptors. *22. Workshop Computerbildanalyse in der Landwirtschaft*, At Wernigerode, Volume: 90, 2016.
- Rose, J.C.; Kicherer, A.; Wieland, M.; Klingbeil, L.; Töpfer, R. & Kuhlmann, H. Towards Automated Large-Scale 3D Phenotyping of Vineyards under Field Conditions. *Sensors*, 16, 2136, 2016. (peer-reviewed)
- Kicherer, A.; Herzog, K.; Bendel, N.; Klück, H.-C.; Backhaus, A.; Wieland, M.; Rose, J.C.; Klingbeil, L.; Läbe, T.; Hohl, C.; Petry, W.; Kuhlmann, H.; Seiffert, U. & Töpfer, R. Phenoliner: A New Field Phenotyping Platform for Grapevine Research. *Sensors*, 17, 1625, 2017. (peer-reviewed)

Einige in dieser Arbeit vorgestellten Aspekte und Ergebnisse wurden bereits in diesen Veröffentlichungen niedergeschrieben und sind als solche referenziert. Inhaltlich umfassen sie jedoch nicht das volle Ausmaß dieser Arbeit und stellen methodisch Teilschritte und Teilerkenntnisse dar. Sie sollten daher als Ergänzung und Dokumentation des Erkenntnisgewinns betrachtet werden.

Alle Abbildungen, soweit der Urheber nicht referenziert wurde, wurden vom Autor erstellt.

Notation

Folgende Notationen werden festgelegt:

- Anzahlen werden mit kaligraphischen Buchstaben bezeichnet: $\mathcal{X}$
- Vektoren werden in fetten Kleinbuchstaben indiziert: $\mathbf{x}$
- Matrizen werden in fetten Großbuchstaben indiziert: $\mathbf{X}$
- Schätzungen werden mit einem Dach gekennzeichnet: $\hat{\mathcal{X}}$

1. Einleitung

1.1 Projekt NoViSys

Die vorliegende Arbeit entstand im Kontext des Projektes *Novel viticulture systems for sustainable production and products* (NoViSys) [NoViSys, 2017], im Arbeitspaket 3: *Quantifizierung der Ausdünneneffekte durch maschinelle Ertragsreduktion und Automatische, sensorgestützte Schätzung der Ertragsreduktion*. Das Projekt begann im Februar 2015 und ist für eine Laufzeit von 5 Jahren vorgesehen. NoViSys wird durch das Bildungsministerium für Bildung und Forschung gefördert (FKZ: 031A349). Folgende Projektpartner sind am Arbeitspaket 3 beteiligt:

- DLR: Dienstleistungszentrum Ländlicher Raum Rheinhessen-Nahe-Hunsrück, Bad Kreuznach
- ERO: ERO Gerätebau GmbH, Niederkumbd
- JKI: Julius-Kühn-Institut für Rebenzüchtung, Siebeldingen
- IGG: Institut für Geodäsie und Geoinformation, Rheinische Friedrich-Wilhelms-Universität Bonn

Die Züchtung sog. *Pilzwiderstandsfähiger Sorten (PIWI)* ist eine der Innovationen im Weinbau, durch die ein ökologischer und ökonomischer Weinbau gefördert werden soll. Etwa ein Drittel aller in Deutschland eingesetzten Fungizide wird im Weinbau verwendet, obwohl die im Weinbau genutzte Fläche lediglich ungefähr 1% der gesamten Agrarfläche Deutschlands entspricht [Kicherer, 2015; Töpfer et al., 2011]. Eine Weiterentwicklung und großflächige Einführung der PIWIs als Alternative zu traditionellen Sorten ist daher von elementarer Bedeutung für die Entwicklung einer ökologischeren und ökonomischeren Landwirtschaft [Töpfer et al., 2011]. PIWIs wurden bislang im *Traditionellen Spaliersystem (TS)* angebaut. Zur Förderung und Evaluierung der PIWIs soll in NoViSys ein neues Erziehungssystem für Weinreben, der *Minimalschnitt im Spalier (MSS)*, untersucht und mit dem traditionellen Spaliersystem in Vergleich gestellt werden. Der Minimalschnitt im Spalier zeichnet sich durch einen reduzierten Arbeitsaufwand von ca. 70%, eine erhöhte Robustheit der Weinreben gegenüber Krankheitsbefall sowie einer verzögerten Reife der Trauben aus [Walg, 2012a,b,c].

Die synergetische Kombination von PIWIs und dem Minimalschnitt im Spalier verspricht eine erhöhte Robustheit gegenüber verschiedenen Pilzen und Krankheiten, sowie einen reduzierten Arbeitsaufwand und ökologischen Fußabdruck durch Einsparung von Spritz- und Schnittmaßnahmen. Zusätzlich soll die Übertragbarkeit des Minimalschnitts im Spalier auf traditionelle Weinsorten untersucht werden. Übergeordnetes Ziel ist dabei eine hohe Weinqualität bei gleichzeitiger Reduktion der Produktionskosten und Umweltbelastungen durch Maschineneinsatz, Dünger und Pestizide [Matese and Di Gennaro, 2015].

TS- und MSS-Erziehungssysteme

Weinreben sind buschige Nutzpflanzen, die in Anbaureihen kultiviert werden und sich typischerweise über mehrere Hektar mit mehreren tausend Pflanzen erstrecken, dem sog. *Wingert*. Im Folgenden werden die jeweilige Charakteristik des traditionellen Spaliersystems und des Minimalschnitts im Spalier erläutert.

Es wird jeweils von *TS- bzw. MSS-Anbaureihen* bzw. dem TS- bzw. MSS-Erziehungssystem gesprochen. Beide Erziehungssysteme unterscheiden sich hinsichtlich der Beschaffenheit des Blattwerks, der Trauben sowieso der Verteilung der Trauben an der Weinrebe deutlich voneinander. Abbildung 1.1 zeigt die zwei Erziehungssysteme vergleichend nebeneinander in einer typischen Feldsituation.

Im MSS Erziehungssystem weisen die Weinreben ein sehr dichtes Blattwerk bei kleineren Blattflächen auf. Das Blattwerk besitzt eine Tiefe von 60 – 100cm bei einer Ausdehnung des Blattwerks von 80 – 210cm über dem Boden. Die Trauben sind weitläufig über die gesamte Höhe und Tiefe des Blattwerks verteilt, wachsen

daher nur selten dicht beieinander und sind im Vergleich zur TS-Erziehung von geringerem Volumen. Der Anteil der Trauben, die für einen im Korridor befindlichen Betrachter durch Blätter und Trauben verdeckt sind, im Weiteren *Verdeckungsgrad* genannt, ist durch das dichte und tiefe Blattwerk sehr hoch. Die Trauben selbst bestehen aus vergleichsweise weniger und kleineren Beeren, die in geringer Dichte und locker am Stielgerüst wachsen.

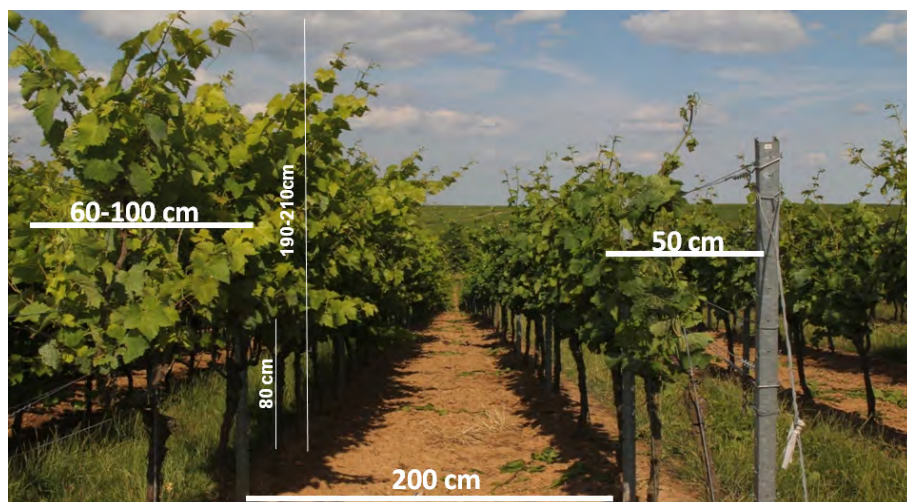


Abbildung 1.1: Zwei Anbaureihen und ein dazwischen liegender Korridor von 2m. Links der Minimalschnitt im Spalier (MSS), rechts der traditionelle Spalieranbau (TS), aus Rose et al. [2016].

Abbildung 1.2 skizziert die unterschiedlich Charakteristik beider Erziehungssysteme hinsichtlich des Blattwerks, der Größe der Trauben und Beeren, sowie der Verteilung des Traubenbestandes.

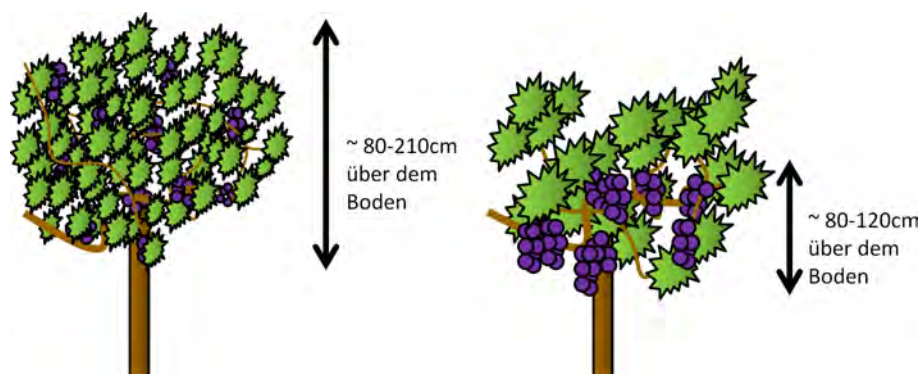


Abbildung 1.2: Die Morphologie des MSS- und TS-Erziehungssystems im Vergleich. Die Erziehungssysteme unterscheiden sich in Blattgröße, Dichte des Blattwerks, sowie Größe und Verteilung der Trauben und Beeren.

Im TS-Erziehungssystem zeichnen sich die Weinreben durch ein luftiges Blattwerk mit vergleichsweise größeren Blattflächen aus. Die Tiefe des Blattwerks beträgt rund 50cm, bei einer Ausdehnung des Blattwerks von 80 – 170cm über dem Boden. Trauben befinden sich vornehmlich in einem Abstand von 80 – 120cm vom Boden, in höheren Zonen des Blattwerks befinden sich nur gelegentlich Trauben. Im Vergleich zur MSS-Erziehung sind die Trauben voluminöser und wachsen im geringen Abstand weniger Zenti- oder Millimeter zueinander. Dadurch sind Trauben 1. häufiger durch andere Trauben verdeckt und 2. räumlich schwieriger voneinander unterscheidbar. Hingegen ist der Verdeckungsgrad durch Blätter durch das luftige Blattwerk vergleichsweise geringer. Die Trauben selbst bestehen aus durchschnittlich mehr und größeren Beeren, die dicht und kompakt zusammen am Stielgerüst wachsen.

Um TS- und MSS-Erziehungssysteme vergleichen zu können, werden in NoViSys sog. **phänotypische Parameter** erhoben, welche die Merkmale der Pflanze quantifizieren, bspw. Blattfläche, Klima im Blattwerk, Größe und Anzahl der Früchte, Chlorophyllgehalt und Krankheitsbefall. Die Gesamtheit der Merkmale einer Pflanze wird dabei als sog. *Phänotyp* bezeichnet, die Erhebung und Bewertung dieser phänotypischen Parameter als **Phänotypisierung**. Im Arbeitspaket 3 von NoViSys soll der **Ertrag** als wichtiger phänotypischer Parameter bestimmt werden. Gemeint ist das tatsächlich geerntete Gewicht des Traubenbestandes, welches sich aus der **Gesamtanzahl der vorhandenen Trauben und Beeren, sowie der Beerengröße** bestimmt [Kicherer, 2015; Nuske et al., 2014b]. Diese drei Größen werden als sog. **Ertragsparameter** bezeichnet.

Ertrag, Ertragsparameter und Ertragsschätzung

Bereits gezählte Teilanzahlen der Ertragsparameter sind mit dem Ertrag korreliert [Aquino et al., 2018] und können als Indikatoren des zu erwartenden Ertrags fungieren [Kicherer, 2015]. Nach Clingeffer et al. [2001] ergibt sich die Varianz des Ertrages zu 60% aus der Varianz der Traubenanzahl, zu 30% aus der Varianz der Beerenanzahl pro Traube und zu 10% aus der Varianz der Beerengröße. Die Beerengröße dient zusätzlich als wichtiger Indikator für aromabezogene Merkmale wie den Zuckergehalt [Matthews and Nuzzo, 2007]. Für die Untersuchungen in NoViSys ist es aus mehreren Gründen notwendig, die Ertragsparameter und den Ertrag bereits vor der Ernte der Trauben am Ende der Saison abschätzen zu können:

- **Ertragsreduktion:** Die Anzahl der Trauben und Beeren an einer Weinrebe ist entscheidend für die Weinqualität [Concurso et al., 2016; Naor et al., 2002]. Weinreben mit weniger Trauben erzeugen Wein von höherer Qualität, bspw. durch eine erhöhte Sonneneinstrahlung pro Traube und eine erhöhte Zuckerzufuhr pro Beere. Sowohl in TS- wie in MSS-Anbaureihen werden daher Trauben zur Optimierung der Weinqualität entfernt. Man spricht von einer *Ausdünnung* des Traubenbestandes. Um eine optimale Weinqualität zu gewährleisten, darf die Anzahl der nach der Ausdünnung an der Weinrebe verbleibenden Trauben weder zu hoch, noch zu gering ausfallen. Um zielgerichtet Trauben zur Optimierung der Weinqualität entfernen zu können, ist es daher wichtig, Kenntnis über die Anzahl der vorhandenen Trauben und den zu erwartenden Ertrag noch vor dem Ausdünnen zu besitzen.
- **Vergleichende Untersuchungen:** Die Anzahl der Trauben und Beeren, sowie der zu erwartende Ertrag von TS- und MSS-Anbaureihen soll zu zwei Reifestadien der Weinreben verglichen werden, um tiefere Kenntnisse über die Entwicklung der Trauben und Beeren in beiden Erziehungssystemen zu erlangen und Traubenverluste während der Saison zu quantifizieren. Um die Entwicklung ausgesuchter Anbaureihen über die Saison hinweg verfolgen zu können, dürfen die Trauben daher nicht vorzeitig geerntet werden.

Zu betrachtende Reifestadien

Winzer bewerten das Reifestadium von Weinreben gemäß der BBCH-Skala (Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie [Lorenz et al., 1995]). Diese Skala beschreibt den Entwicklungsgrad phänotypischer Parameter auf Ebene der Pflanzenorgane, wie die Anzahl der Trauben, Erscheinungsbild der Blüten und des Holzes, den Grad der Blattentfaltung oder die Größe und Farbe einzelner Beeren. Die einzelnen Reifestadien werden über sequentiell steigende Werte von 0 – 97 indiziert [Kicherer, 2015]. Die Bewertung der Entwicklung der Weinreben beider Erziehungssysteme soll in NoViSys anhand des Zustandes in den zwei Reifestadien BBCH75 und BBCH89 stattfinden. Im Folgenden wird die Morphologie der Weinreben in beiden Reifestadien näher beschrieben und die Bedeutung der Reifestadien für NoViSys erläutert.

Die Ausdünnung des Traubenbestandes findet im BBCH75 Reifestadium statt. Anbaureihen im MSS-Erziehungssystem neigen zu starker Überproduktion von Trauben. Der Traubenbestand von Anbaureihen des MSS-Erziehungssystems wird daher maschinell durch einen Erntetraktor durch Schütteln an den Weinreben reduziert. Die Anzahl entfernter Trauben variiert dabei mit der eingestellten Schüttelintensität des Erntetraktors. Im TS-System wird die Ertragsreduktion auf händische Weise mit einer Schere individuell für einzelne Trauben vorgenommen.

Um die optimale Anzahl von Trauben zu entfernen, müssen die Ertragsparameter und der Ertrag für beide Erziehungssysteme bereits vor der Ausdünnung abgeschätzt werden. Zur Quantifizierung der Ausdünnungsmaßnahmen muss der Ertrag nach der Ausdünnung erneut bestimmt werden, um eventuell weitere Ausdünnungen vorzunehmen.



A



B

Abbildung 1.3: Typische Morphologie der MSS- und TS-Erziehung im BBCH75-Reifestadium. **A:** Eine Riesling-Anbaureihe nach der MSS-Erziehung. **B:** Eine Chardonnay-Anbaureihe nach der TS-Erziehung.

Abbildung 1.3 zeigt jeweils eine TS- und MSS-Anbaureihe im BBCH75 Reifestadium. Die morphologischen Unterschiede in den Erziehungssystemen bzgl. Blattdichte oder Traubenverteilung sind in diesem Reifestadium bereits ausgeprägt. Beeren im BBCH75 Reifestadium sind, unabhängig von der Sorte, ausnahmslos von hellgrüner Färbung. Abbildung 1.3 (A) zeigt eine Riesling-Anbaureihe nach der MSS-Erziehung im BBCH75 Reifestadium. Die Beeren besitzen einen Durchmesser von 6 – 10mm. Abbildung 1.3 (B) zeigt eine Chardonnay-Anbaureihe nach der TS-Erziehung im selben Reifestadium. Die Beeren sind hier bereits etwas größer als die Beeren des MSS-Erziehungssystems und besitzen einen Durchmesser von 7 – 11mm.



A



B

Abbildung 1.4: Zwei Riesling Anbaureihen im BBCH89 Reifestadium, aus Rose et al. [2016]. **A:** MSS-Erziehung. **B:** TS-Erziehung.

Um die Entwicklung der Trauben und des Ertrages über die Saison [Chabbi and Loescher, 2017; Furbank and Tester, 2011; Kicherer, 2015] für beide Erziehungssysteme bewerten zu können, soll eine weitere Schätzung der Ertragsparameter und des Ertrages für beide Erziehungssysteme im BBCH89 Reifestadium stattfinden. In diesem Reifestadium sind die Trauben kurz vor der Erntereife. Die Betrachtung in diesem späten Reifestadium ist nötig, da im Laufe der Saison Trauben und Beeren bedingt durch Umweltbedingungen, Krankheiten und Pilzbefall oder durch Verletzungen beim Schütteln nachträglich abfallen können. Dieser Verlust soll ebenfalls quantifiziert und mit den Schätzungen im BBCH75 Reifestadium verglichen werden.

Abbildung 1.4 (A) zeigt eine Riesling Anbaureihe nach der MSS-Erziehung im Reifestadium BBCH89. Die Farbe der Beeren ist mattgrün, teilweise bräunlich. Ihr Durchmesser beträgt in diesem Stadium $6\text{--}14\text{mm}$. Abbildung 1.4 (B) zeigt eine TS-Anbaureihe im BBCH89 Reifestadium. Die Farbe der Beeren ist auch hier mattgrün mit bräunlichem Einschlag. Der Durchmesser der Beeren ist im Vergleich zum MSS-Erziehungssystem größer und liegt bei $6\text{--}18\text{mm}$.

Traditionelle und moderne Methoden der Ertragsschätzung

Bisher geschah die Schätzung der Ertragsparameter und des Ertrages auf manuelle Weise durch einen erfahrenen Winzer. Bei diesen traditionellen Methoden der Ertragsschätzung werden üblicherweise stichprobenartig Teilanzahlen der Trauben und Beeren visuell gezählt, vermessen und basierend auf Mittelwerten aus zurückliegenden Saisons und der Erfahrung des Winzers der zu erntende Ertrag über mehrere Anbaureihen extrapoliert [Nuske et al., 2014b]. Eine etwas genauere Methode ist die Zählung und Gewichtsbestimmung weniger Trauben und Beeren entweder direkt im Feld oder deren Ernte und anschließende Zählung und Gewichtsbestimmung. Die Ergebnisse werden anschließend extrapoliert, um eine Schätzung des zu erwartenden Ertrages zu erhalten. Diese traditionellen Methoden der Schätzung der Ertragsparameter und des Ertrages sind ungenau, langsam, arbeits- und kostenintensiv, nur stichprobenartig, subjektiv und aufgrund der Variabilität von Weinreben durch Stichprobenfehler verzerrt.

Für die Untersuchungen in NoViSys sollen die Ertragsparameter mehrerer Kilometer von Anbaureihen möglichst genau, vollständig, zu gleichen Zeitpunkten, wiederholt in den BBCH75 und BBCH89 Reifestadien und über mehrere Saisons hinweg bestimmt werden. Dies hat folgende Gründe:

- Die Verlässlichkeit der Ertragsschätzung steigt, je vollständiger die Ertragsparameter bestimmt und je mehr Weinreben für die Ertragsschätzung herangezogen werden. Eine vollständige Zählung aller Trauben und Beeren und die Bestimmung der Größe jeder Beere im Wingert würde somit eine präzise Schätzung des Ertrages noch vor der Ernte ermöglichen. Physikalisch ist es aber aufgrund der Größe des Wingert und der großen Anzahl von Trauben sehr aufwändig, die Ertragsparameter auf manuelle Weise vollständig und für alle Weinreben des Wingert zu bestimmen.
- Um Vergleichbarkeit der Ergebnisse zu gewährleisten sollten alle Messungen möglichst zu gleichen Zeitpunkten und damit identischen Umweltbedingungen stattfinden. Um eine hohe Datenmenge zu erreichen, müssen dabei mehrere Kilometer von Weinreben untersucht werden.
- Die Bestimmung des Ertrages muss des Weiteren nicht-invasiv geschehen, um den vollständigen Traubenbestand zu erhalten und die Entwicklung der Pflanzen in den zwei Reifestadien verfolgen zu können.

Mit den traditionellen Methoden können diese Ansprüche jedoch nicht erfüllt werden. In NoViSys sollen die Einschränkungen in Genauigkeit, Geschwindigkeit, Objektivität und Kosten, denen die Schätzung der Ertragsparameter und des Ertrages durch die traditionellen Methoden unterlegen ist, durch den Einsatz nicht-invasiver, sensorgestützter Methoden überwunden werden. Dabei sollen automatisch objektive, genaue, zeitlich und örtlich hoch aufgelöste Zählungen und Vermessungen der Trauben und Beeren und eine Schätzung des Ertrages von einer Vielzahl von Weinreben vorgenommen werden [Herzog et al., 2014; Kicherer, 2015; Whalley and Shanmuganathan, 2013].

1.2 Aufgabe

Aus den oben besprochenen Vorhaben in NoViSys ergibt sich die Aufgabe dieser Arbeit:

Entwicklung einer nicht-invasiven, sensorgestützten Hochdurchsatz-Methode zur automatischen Schätzung der Ertragsparameter und eine auf den Ertragsparametern basierende Schätzung des Ertrages für TS- und MSS-Erziehungssysteme im Reifestadium BBCH75 und BBCH89 unter Feldbedingungen.

Eine für diese Aufgabe konzipierte Methode muss auf die Morphologie des jeweiligen Erziehungssystems und Reifestadiums angepasst sein und den Anforderungen von NoViSys hinsichtlich Kosten, Genauigkeit, Objektivität und Geschwindigkeit gerecht werden. Die Methode muss daher unten aufgeführte Kriterien erfüllen:

1. Nicht-Invasiv: Blattwerk und Traubenbestand müssen für die Untersuchungen bis zur Ernte intakt bleiben. Da Trauben und Beeren direkt gezählt und vermessen werden sollen, bietet sich die Verwendung einer **optischen Sensorik** an.
2. Nahbereich: Die Weinreben sind nur aus den Korridoren zwischen zwei Anbaureihen einsehbar. Die Erfassung der Trauben muss daher aus dem $2m$ breiten Korridor zwischen den Anbaureihen heraus erfolgen. Dies bedingt den Einsatz von optischen Sensoren, die für den Einsatz im **Nahbereich** geeignet sind.
3. Hochaufgelöst: Beeren stellen den kleinsten möglichen Gewichtsanteil des gesamten Ertrages dar. Je präziser und vollständiger sie gezählt werden können, desto präziser wird die Ertragsschätzung. Im BBCH75 Reifestadium besitzen Beeren einen Durchmesser von nur wenigen Millimetern, mit maximal 18mm im BBCH89 Reifestadium. Die Daten müssen daher einen Detailgrad besitzen, der die **direkte Zählung und Vermessung einzelner Beeren** im BBCH75 und BBCH89 Reifestadium ermöglicht.
4. Maßstab: Die Daten sollen einen **metrischen Maßstab** zur Vermessung der Trauben und Beeren besitzen.
5. Hochdurchsatz: Um eine ausreichende Datenmenge und einen identischen Zeitbezug für vergleichende Untersuchungen zu erhalten, sollen mindestens drei Kilometer von Anbaureihen an einem Arbeitstag erfasst werden können. Die Datenerhebung soll daher **schnell und aus der Bewegung** geschehen.
6. Automation: Die Datenerhebung sowie die Schätzung der Ertragsparameter und des Ertrages aus den Sensordaten sollen **automatisch** erfolgen, um den Kosten- und Personalaufwand zu reduzieren. Für die Datenerhebung bedingt dies den Einsatz einer **mobilen Sensorplattform**, welche die Datenerhebung automatisch durchführt, sowie die Entwicklung von Algorithmen zur automatischen Auswertung der Daten.
7. Standardisierte Aufnahmebedingungen: Um Reaktionen der Weinreben auf Umweltphänomene zeitnah und zeitlich hoch aufgelöst quantifizieren zu können, müssen Datenerhebungen mehrmals in der Saison und unabhängig von äußeren Beleuchtungsverhältnissen und Wetterbedingungen möglich sein. Wechselnde Beleuchtungsverhältnisse oder Regen resultieren in heterogenen Daten, was die Datenauswertung erschweren kann. Eine gleichzeitige Abschirmung von Sensor und Weinreben unter Nutzung künstlicher Beleuchtungsquellen schafft **standardisierte Aufnahmebedingungen** auch unter Feldbedingungen, welche die Einsatzmöglichkeiten der Sensorplattform erweitert und standardisierte Daten für die Auswertung generiert. Dies bedingt notwendig die Verwendung einer entsprechend **großen mobilen Sensorplattform**, welche sowohl Weinreben wie Sensoren gleichzeitig umschließen kann.
8. Vollständigkeit: In beiden Erziehungssystemen sind die Trauben durch Äste, Blätter oder andere Trauben verdeckt, wobei insbesondere das MSS-Erziehungssystem aufgrund seines dichten Blattwerks einen sehr hohen Verdeckungsgrad der Trauben aufweist. Der Traubenbestand einer Anbaureihe soll dennoch **möglichst vollständig** erfasst werden, um die Variabilität der Weinreben zu berücksichtigen und den

Einfluss von Stichprobenfehlern zu reduzieren. Der Sensor sollte daher mehrere Perspektiven entlang der Horizontalen und der Vertikalen auf die Anbaureihe einnehmen, 1. um **Verdeckungen** der Trauben durch Äste, Blätter oder andere Trauben weit möglichst zu **kompensieren** und 2. um die über die **gesamte Höhe** des Blattwerks verteilten Trauben des MSS-Erziehungssystems zu erfassen.

9. Zeitnahe Referenzwerte: Trotz der im vorherigen Punkt ausgeführten Maßnahme, mit der Verdeckungen des Traubenbestandes möglichst verringert werden soll, ist es nicht möglich, alle Beeren und Trauben mittels einer mobilen Sensorplattform vollständig zu erfassen. Zwangsläufig wird nur eine Teilanzahl der tatsächlich vorhandenen Trauben und Beeren mit dem optischen Sensor beobachtet werden können, die als *sichtbare Ertragsparameter* bezeichnet werden. Dies resultiert 1. aus der einschränkenden Positionierung des Sensors in den Korridoren und des Wachsens der Beeren in Trauben. Dadurch kann nur die dem Sensor zugewandte Seite der Trauben erfasst werden. Beeren auf der dem Sensor abgewandten Seite der Traube, sowie im Inneren der Traube werden nicht erfasst. Eine gewisse Anzahl an Trauben in der Tiefe des Blattwerks wird 2. stets verdeckt sein und vom Sensor niemals erfasst werden können. Diese **nicht detektierbaren Teilanzahlen** sind durch **Kalibrierwerte** zu berücksichtigen, mit denen die sichtbaren Ertragsparameter korrigiert werden. Um eine realistische Schätzung des aktuellen Bestandes zu erhalten, sollten diese jedoch nicht historisch sein, sondern aus **zeitnah erhobenen Referenzmessungen** der aktuellen Saison abgeleitet werden.
10. Lokalisierung: Um die Daten räumlich interpretieren und in einem Geoinformationssystem (GIS) verwalten zu können, müssen die Daten **georeferenziert** sein. Für den Einsatz im offen liegenden Wingert bieten sich dafür GPS-Sensoren an.

Unter Bezug auf die gestellte Aufgabe und unter Berücksichtigung der oben genannten Kriterien ergeben sich folgende zu lösende **Teilaufgaben zur Schätzung der Ertragsparameter und des Ertrages**:

1. **Datenerhebung:** Entwicklung und Untersuchung *mobiler Sensorplattformen* für die Datenerhebung im Hochdurchsatz unter Feldbedingungen für beide Erziehungssysteme. Die Teilaufgabe umfasst die Untersuchung einer bereits existierenden Sensorplattform, sowie die Realisierung und Evaluierung einer innovativen neuen Sensorplattform, welche eigens für die Belange von NoViSys entwickelt wurde, sowie die Vorverarbeitung und Bewertung der erhobenen Daten.
2. **Datenklassifikation:** Entwicklung einer geeigneten algorithmenbasierten *Klassifikationsstrategie*, um in den Daten Beeren und Trauben von Blättern, Ästen und anderen Bestandteilen automatisch zu unterscheiden.
3. **Quantifikation:** Entwicklung einer algorithmenbasierten Strategie zur automatischen *Schätzung der Ertragsparameter und des Ertrages* aus den klassifizierten Daten.

1.3 Aufbau der Arbeit

- Kapitel 2 bespricht die bisherigen Forschungsbeiträge, die sich mit der Schätzung der Ertragsparameter und des Ertrages von Weinreben befassen. Das Kapitel ordnet die Beiträge außerdem hinsichtlich der Erfüllung der oben definierten Kriterien ein. Zuletzt wird der Beitrag der vorliegenden Arbeit mit Hinsicht auf diese Kriterien formuliert.
- Kapitel 3 enthält die Erläuterungen der umfangreichsten und wichtigsten Methoden, die in dieser Arbeit zum Tragen kommen.
- Kapitel 4 schildert Voruntersuchungen zur Datenerhebung mit der bereits existierenden Sensorplattform, zur optimalen Klassifikation der Anbaureihen und beschreibt die Auswertemethode zur Detektion der Beeren und Trauben in den Punktwolken. Die Genauigkeit der Auswertemethode wird für beide Reifestadien und Erziehungssysteme anhand von Referenzwerten untersucht.

- Kapitel 5 stellt die für NoViSys neu entwickelte Sensorplattform vor, beschreibt die mit ihre vorgenommene Datenerhebung, die Klassifikation der Punktwolken und die Berechnung von Kalibrierwerten, mit denen die Gesamtanzahl aller Beeren und Trauben geschätzt werden kann. Basierend auf den geschätzten Ertragsparametern wird eine Schätzung des Ertrages für beide Erziehungssysteme im BBCH89 Reifestadium vorgenommen und anhand von Referenzwerten evaluiert.
- Kapitel 6 fasst die wichtigsten Ergebnisse nochmal zusammen und formuliert ein Fazit hinsichtlich der oben gestellten und in dieser Arbeit zu lösenden Aufgabe. Ein Ausblick, in dem Verbesserungsmöglichkeiten, künftige lohnenswerte Untersuchungen sowie weitere Einsatzmöglichkeiten der entwickelten Sensorplattform und Auswertemethode abseits von NoViSys beschrieben werden, beschließt die Arbeit.

2. Stand der Forschung

Heutzutage sind bereits eine Vielzahl von Methoden zur Schätzung der Ertragsparameter und des Ertrages mittels optischer Sensoren entwickelt worden [Herzog et al., 2014; Kicherer, 2015] (für weitere phänotypische Parameter sei bspw. auf Whalley and Shanmuganathan [2013] verwiesen). Dabei wurden primär bildbasierte Methoden entwickelt, neuere Ansätze erforschen außerdem die Potentiale von 3D-Methoden wie Laserscannern, (Multi-View)-Stereo Ansätzen oder Structured-Light Scannern. Im Folgenden werden diese Forschungsbeiträge eingehender besprochen und eingeordnet.

2.1 Bildbasierte 2D-Methoden

Der Großteil der bisherigen Methoden basiert auf Bildverarbeitungsmethoden unter Nutzung von RGB- oder Multispektral-Bildern. Dabei werden sowohl Beeren und Trauben gezählt und vermessen. Zur Zählung der Beeren wird häufig die zirkuläre Hough-Transformation [Duda and Hart, 1972] verwendet.

Diago et al. [2015] bestimmen im Labor die Anzahl und das Gewicht von Beeren, sowie das Gewicht von Trauben aus RGB-Bildern. Die zirkuläre Hough-Transformation wird zur Kreisdetektion genutzt, wobei detektierte Kreise als Beeren gelten. Referenzmessungen stellen einen funktionalen Zusammenhang zwischen geschätztem Kreisradius und tatsächlichem Beerengewicht her. Abbildung 2.1 illustriert die Vorgehensweise im Labor.

Liu et al. [2015] zählen Bilder aus einem einzigen RGB-Bild einer Traube, welches vor einem künstlichen Hintergrund aufgenommen wurde. Eine Schwellwertmethode trennt die Traube vom Hintergrund. An den Randgebieten der Traube werden mittels der zirkulären Hough-Transformation initiale Beerenkandidaten detektiert, von denen aus das Bild iterativ von oben nach unten nach weiteren Kreisen durchsucht wird. Mittels eines Kalibrierwertes, der die Kompaktheit der Traube berücksichtigt, wird die Beerenanzahl geschätzt.



Abbildung 2.1: Einsatz des stationären Sensors aus Diago et al. [2015] im Labor.

Der. sog. *PHENObot*, eine mobile Sensorplattform mit Kettenlaufwerk [Kicherer et al., 2015a], steuert im Feld automatisiert vorab bestimmte GPS-Positionen einzelner Weinreben an und nimmt bei Nacht autonom georeferenzierte, ausgeleuchtete RGB-Bilder auf. Vorab bestimmte GPS-Wegpunkte ermöglichen den autonomen Übergang des PHENObots zwischen den Anbaureihen. In den Bildern werden anschließend

mit einem Algorithmus namens *BIVcolor* [Roscher et al., 2014] automatisch die Größe der Beeren durch zirkuläre Hough-Transformation, sowie die Farbe der Beeren bestimmt und durch die Georeferenzierung der zugehörigen Weinrebe in der Anbaureihe zugeordnet.

Andere Methoden nutzen Ellipsenfitting zur Modellierung der Beeren, um die nicht vollständig runde Geometrie der Beeren zu berücksichtigen.

Das Berry Analysis Tool (BAT) [Kicherer et al., 2015b] besteht aus einer Digitalkamera und einer 300x300mm messenden, perforierten Metallplatte, auf welcher Beeren platziert werden. Mittels Schwellwertmethoden wird der Hintergrund vom Vordergrund getrennt, die Maße der Platte werden als Maßstab für die Umrechnung von Pixeleinheiten in metrische Einheiten verwendet. Anschließend werden Pixelklassifikation, morphologische Operationen und Ellipsenfitting verwendet, um Anzahl und Größe der Beeren zu bestimmen.

Weitere bildbasierte Methoden nutzen Form, Farbe und die Lichtreflexion auf der Beerenmitte zur Zählung von Beeren.

In Aquino et al. [2017] wird die Anzahl der Beeren aus einem RGB-Bild über mathematische Operationen und Pixelklassifikation bestimmt. Trauben werden vor einem schwarzen Hintergrund fotografiert und durchlaufen ein mehrstufiges Evaluationsverfahren, welches die Pixelintensität auf der Beerenmitte sowie weitere Merkmale, Connected Components und überwachte Klassifikationsmethoden umfasst. Das Verhältnis zwischen algorithmisch detektierter und tatsächlicher Beerenanzahl wird zuletzt über Lineare Regression mittels Referenzwerten bestimmt. In Aquino et al. [2018] wird die eben beschriebene Methode auf Bilder angewandt, die von einer Sensorplattform unter Feldbedingungen bei Nacht mit einer RGB-Kamera aufgenommen wurde. Die Sensorplattform nutzt GPS für die Georeferenzierung der Bilder und das Zusammenfügen von stets drei Bildern zu einem einzigen Bild, um Mehrfachzählungen in redundanten Bildinhalten zu vermeiden. Vorder- und Hintergrund werden anhand der nachtschwarzen Pixel des Hintergrunds voneinander unterschieden.



Abbildung 2.2: Die Sensorplattform aus Nuske et al. [2014b] im Feldeinsatz bei Nacht (modifiziert).

Neben den Beeren werden ebenfalls Trauben vermessen und gezählt. Das CAT (Cluster Analysis Tool) [Kicherer et al., 2014] nutzt RGB-Bilder ganzer Trauben, um Länge, Breite und Kompaktheit der Trauben zu messen. Trauben werden geerntet und vor dem Sensor positioniert. Ein im Bild angebrachter Maßstab ermöglicht die Maßstabumrechnung von Pixeleinheiten in metrische Einheiten. Das Bild wird in Hintergrund und Traube klassifiziert. Ebenfalls mittels zirkulärer Hough-Transformation sowie den Reflexionspunkte des

Kamerablitzes auf der Beerenmitte wird die Anzahl der Beeren geschätzt. Die Dimension der Boundingbox der Traube gelten als Traubenmaße. Die Traubenfläche wird anschließend mit der Breite und Länge der Boundingbox verrechnet, um die Kompaktheit der Traube zu bestimmen.

Nuske et al. [2014b] nehmen unter Feldbedingungen RGB-Bilder automatisch von einer auf einem Erntetraktor installierten Sensorplattform aus der Bewegung heraus auf. Textur, Geometrie und die Reflexionen des Lichtblitzes der Beleuchtungseinheit auf der Beerenmitte werden genutzt, um mit einem randomisierten KD-forest [Breiman, 2001] Beeren zu detektieren und zu zählen. Eine Stereokamera bestimmt zusätzlich die Position des Erntetraktors relativ zur Anbaureihe. Die Position wird in einer heuristischen Methode dazu benutzt, Mehrfachzählungen der Beeren bei überlappenden Bildinhalten zu verhindern. Benachbarte Beeren werden zu Clustern zusammengefasst und Cluster unterhalb einer Schwellwertflächengröße gelöscht. Die verbleibenden Cluster werden zunächst als Trauben aufgefasst. Jedoch werden keine Ergebnisse für die Traubenmenge vorgestellt, da sich überlappende Trauben nicht erfolgreich unterscheiden ließen. Kalibrierwerte aus Referenzmessungen korrigieren die algorithmisch detektierte Anzahl der sichtbaren Beeren um die Anzahl nicht sichtbarer Beeren. Basierend auf der korrigierten Beerenanzahl wird der Ertrag geschätzt. Abbildung 2.2 zeigt den Einsatz der Sensorplattform unter Feldbedingungen bei Nacht.

In Mirbod et al. [2016] wird auf die Algorithmen in Nuske et al. [2014b] zur Zählung und zur Bestimmung der Größe von Beeren zurückgegriffen. Eine kalibrierte Stereokamera auf einer bewegten Plattform nimmt RGB-Bilder aus der Fahrt heraus auf. Die bekannte Länge der Basislinie der Stereokamera ermöglicht die Überführung der geschätzten Beerenradii von Pixeleinheiten in metrische Einheiten. Kalibrierwerte für die Beerendichte innerhalb einer Traube und den Verdeckungsgrad der nicht-sichtbaren Beeren werden mit der Beerenmenge und -größe verrechnet und eine Schätzung des Ertrages erhalten.

Liu and Whitty [2015] nehmen mit einer konventionellen Farbkamera aus der Hand RGB-Bilder von Weinreben unter Tageslichtverhältnissen im Feld auf. Schwellwertdiskriminierung im HSV-Raum (Hue-Saturation-Value) wird anschließend verwendet, um in den Bildern Regionen potentieller Traubenpixel zu identifizieren. Für jede potentielle Region von Traubenpixeln wird die Boundingbox berechnet und so eine erste Zählung der Trauben erhalten. Um in den potentiellen Traubenpixelregionen Nicht-Traubenregionen zu erkennen und zu entfernen, wird eine Support Vector Machine [Cortes and Vapnik, 1995], basierend auf Deskriptoren wie Farbe, Kompaktheit, Höhe im Bild oder Solidität, genutzt. Die verbleibenden Anzahl der Boundingboxes der Traubenregionen gilt als Schätzung der Traubenanzahl.

In Luo et al. [2016] werden RGB-Bilder von Hand unter variierenden Beleuchtungsverhältnissen aufgenommen und mehrere lineare Schwellwertklassifikatoren basierend auf Pixelfarben erstellt. Um Traubenpixel zu identifizieren, werden diese in einem weiteren Schritt zu einem einzigen Klassifikator kombiniert. Für alle Regionen von Traubenpixeln wird ihre Boundingbox berechnet. Kleine Regionen unterhalb einer Schwellwertgröße gelten als Fehlklassifikationen und werden gelöscht. Die Anzahl der verbliebenen Boundingboxes gilt als Schätzung der Traubenanzahl.

Andere Methoden weichen von der semantischen Modellierung anhand von Kreisen, Ellipsen oder Boundingboxen ab und verwenden allein Pixelflächen zur Bestimmung von Gewicht und Anzahl von Beeren/Trauben. Diago et al. [2012] klassifizieren Pixel von RGB-Bildern einer Anbaureihe in die Klassen Jung-/Altblätter, Holz, Trauben und künstlicher Hintergrund. Trainingsdaten aus Pixelfarbwerten werden genutzt, um die Mahalanobis Distanz [Mahalanobis, 1936] zur Unterscheidung von Farbgruppen zu berechnen. Pixelflächen der Klassen werden anschließend genutzt, um eine Schätzung der Blattfläche und des Traubengewichts zu treffen.

Rabatel and Guizard [2007] nutzen einen orangefarbenen Hintergrund, an welchem eine Digitalkamera mit Blitzfunktion angebracht ist. Eine Traube wird auf dem Hintergrund platziert und fotografiert. Watershed-Algorithmen und Baumstrukturen für zusammenhängende Pixel werden anschließend angewandt, um die Beerengröße zu bestimmen. Abbildung 2.3 illustriert den Einsatz des tragbaren Sensors im Feld. Die Arbeit von Fernández et al. [2013] nutzt statt des RGB-Raums den Multispektral-Raum. Darin wird eine Kamera und ein motorisiertes Filterrad zur Aufnahme multispektraler Bilder verwendet. Um die Bildinhalte in die Klassen Laub, Äste, Stängel, Hintergrund und Frucht zu klassifizieren, werden die verschiedenen spektralen Signaturen der Bild in einem sequentiellen k-means-clustering [Lloyd, 1982] Ansatz genutzt.



Abbildung 2.3: Der tragbare Sensor aus Rabatel and Guizard [2007] mit künstlichem Hintergrund im Feldeinsatz (modifiziert).

2.2 3D-Methoden

Einige Arbeiten nutzen 3D-Methoden zur Rekonstruktion der Geometrie von Beeren und Trauben. Dabei werden die Beeren zumeist durch Kugelapproximationen modelliert.

In Ivorra et al. [2015] werden ein Stereokamera paar und 3D-Deskriptoren zur Erstellung eines 3D-Modells einer Traube genutzt. Anhand des 3D-Modells werden anschließend Anzahl, Gewicht und Größe der Beeren, sowie das Volumen und die Kompaktheit der Traube bestimmt.

Schöler and Steinhage [2015] konstruieren ein vollständiges Modell der Traubenarchitektur aus hochauflösenden Laserscanpunktwolken. Sie implementieren einen dreistufigen Rekonstruktionsalgorithmus, der zunächst datenbasiert Kugeln als Beerenkandidaten aus der Punktwolke schätzt. Im zweiten Schritt werden basierend auf den Beerenkandidaten modellbasierte Schätzungen der Beeren und Stängel im nicht einsehbaren inneren Teil der Traube getroffen. Eine Optimierung der daten- und modellbasierten Schätzungen erfolgt im dritten Schritt.

Mack et al. [2017a] klassifizieren eine hochauflösende 360°-Laserscanpunktwolke einer vollständigen Weintraube mit geometrischen Deskriptoren automatisch in Stängel und Beeren. Segmentierungsmethoden und Kugelapproximationen werden anschließend zur Zählung der Beeren verwendet.

Einige 3D-basierte Methoden wurden ebenfalls für den Feldeinsatz untersucht. Dabei wurden sowohl Laserscanner, als auch Structure-from-Motion/Multi-View-Stereo Methoden (SfM/MVS) verwendet.

Mack et al. [2017b] nutzen mehrere geometrische Deskriptoren zur Klassifikation von 3D-Punktwolken von Weinreben, Weintrauben und einer Blume, die unter Feldbedingungen mit einem tragbaren Laserscanner generiert wurden. Die Punktwolken wurden mit allen untersuchten Deskriptoren unter Nutzung einer Support Vector Machine klassifiziert und die Genauigkeit der Klassifikationen miteinander verglichen. Die Fast-Point-Feature-Histograms [Rusu, 2009] erwiesen sich dabei als bester Deskriptor.

Herrero-Huerta et al. [2015] nutzen eine handelsübliche RGB-Kamera, um von Hand Bilder einzelner Weinreben aus drei horizontalen und zwei vertikalen Positionen aufzunehmen. Mittels einer eigenen SfM/MVS-Methode werden farbige Punktwolken der Anbaureihen erzeugt und über bekannte Maßstäbe in den Punktwolken auf metrische Einheiten skaliert. Trauben werden manuell segmentiert und anschließend das Traubenvolumen und -gewicht, sowie die Anzahl der sichtbaren Beeren durch CAD- und Vermaschungstechniken bestimmt. Anhand der konvexen Hülle wird die Anzahl nicht sichtbarer Beeren geschätzt. Mit Referenzwerten wird ein funktionaler Zusammenhang zwischen Traubenvolumen und Ertrag erzeugt.

Santos et al. [2017] nutzen eine Digitalkamera zur Rekonstruktion farbiger 3D-Punktwolken von Weinreben. Sie verwenden anschließend den geometrischen Deskriptor Fast-Point-Feature-Histograms und Farbdeskriptoren zur Klassifikation der Punktwolken. Mittels eines 3D-Connected Components Verfahrens wurden an-

schließlich individuelle Trauben identifiziert, ihr Volumen durch eine konvexe Hülle approximiert und das Volumen mit Referenzdaten für den Ertrag der Trauben verglichen.

Eine handgehaltene Microsoft Kinect (Microsoft Corp., Redmond, USA) wird in Marinello et al. [2016] verwendet, um Punktwolken von Trauben zu erzeugen. Der Einfluss der Distanz, sowie die Positionierung des Sensors relativ zu den Trauben auf die Schätzung von Volumen und Masse der Trauben wurden untersucht.

Andere Arbeiten bestimmen keine Ertragsparameter und beschränken sich auf die Klassifikation der Daten. Klodt et al. [2015] nutzen eine Stereokamera zur Berechnung von Tiefenbildern von Weinreben. Der Hintergrund des Bildes wird durch die Tiefeninformation von der Weinrebe getrennt. Mittels der Farbinformation wird das Bild der Weinrebe weiter in die Klassen Äste, Blätter und Trauben klassifiziert.

Dey et al. [2012] nehmen mit einer handelsüblichen Kamera Sequenzen von RGB-Bildern vollständiger Anbaureihen auf und erzeugen mittels einer MVS-Methode farbige 3D-Punktwolken. Sie nutzen anschließend auf Hauptkomponentenanalyse basierende geometrische Deskriptoren, um jeden Punkt der Punktwolke den Klassen Ast, Blatt oder Traube zuzuordnen.

2.3 Beitrag dieser Arbeit

Grundsätzlich lassen sich die bisherigen Methoden in zwei Kategorien einteilen:

1. Selektive Erfassung isolierter Trauben mittels ruhender oder getragener optischen Sensoren, oftmals unter Verwendung eines künstlichen Hintergrundes.
2. Indiskriminative Erfassung vollständiger Weinreben mit Blattwerk, Trauben und Objekten des Hintergrundes mittels einer mobilen Sensorplattform oder eines getragenen optischen Sensors, entweder aus der Fahrt heraus oder aus dem Stand.

Der Einsatz der Methoden erfolgt dabei entweder unter homogenen Labor- oder heterogenen Feldbedingungen. Tabelle 2.1 gibt einen Überblick über die bisherigen Methoden unterteilt nach Einsatzort und Vorgehensweise bei der Datenerhebung.

Im Folgenden seien die oben dargelegten Kategorien näher besprochen. Dabei erfolgt eine Einordnung der bisherigen Publikationen hinsichtlich der in dieser Arbeit gestellten Aufgabe und der in Kapitel 1 formulierten methodischen Kriterien, die an eine Lösung dieser Aufgabe gestellt werden.

2.3.1 Einordnung der Fachbeiträge

Die für den Einsatz im Labor konzipierten Methoden bedingen stets die vorherige Ernte von Trauben und deren Positionierung im Arbeitsbereich des Sensors. Diese Methoden verletzen die Kriterien der Nicht-Invasivität, Automation und Datenerhebung im Hochdurchsatz und sind daher grundsätzlich für die Belange in NoViSys nicht geeignet. Einige Methoden übertragen die methodischen Ansätze aus den Laboren auf den Einsatz unter Feldbedingungen und vermeiden dabei die invasive Ernte der Trauben durch Nutzung tragbarer, flexibler Sensoren. Bei diesen Methoden werden die Trauben während der Erfassung an der Weinrebe belassen und die einzelnen Trauben **selektiv** von einem Anwender, mitunter händisch vom Blattwerk isoliert, im Arbeitsbereich des Sensors positioniert oder der Sensor wird direkt in eine günstige Aufnahmeposition zur Erfassung der Traube gebracht. Oftmals wird ein künstlicher Hintergrund hinter der Traube positioniert, um den Hintergrund vom Vordergrund durch einfache Farbdiskriminierung unterscheiden zu können. Die Traube wird somit isoliert von anderen Teilen der Weinrebe erfasst. Tragbare Sensoren sind dafür konzipiert, schnelle Ergebnisse für einzelne Trauben auf nicht-invasive Weise zu erzeugen. Abbildung 2.4 zeigt ein solches Vorgehen. Das händische Selektieren und Positionieren der einzelnen Trauben vor dem Sensor und dem künstlichen Hintergrund oder die günstige Positionierung des Sensor vor den Trauben geschieht jedoch zwangsläufig durch einen Anwender und ist daher sehr langsam und die Datenerhebung erfolgt höchstens

	Labor	Feld
Selektive Erfassung	Diago et al. [2015] Mack et al. [2017a] Kicherer et al. [2014] Kicherer et al. [2015b] Ivorra et al. [2015] Schöler and Steinhage [2015]	Aquino et al. [2017] Liu et al. [2015] Rabatel and Guizard [2007] Mack et al. [2017b]
Indiskriminative Erfassung	–	Santos et al. [2017] Diago et al. [2012] Liu and Whitty [2015] Luo et al. [2016] Fernández et al. [2013] Herrero-Huerta et al. [2015] Dey et al. [2012] Marinello et al. [2016] Klodt et al. [2015] Kicherer et al. [2015a] Nuske et al. [2014b] Mirbod et al. [2016] Aquino et al. [2018]

Tabelle 2.1: Einordnung der Fachbeiträge nach den Einsatzorten Labor und Feld, sowie nach Art der Datenerhebung.



Abbildung 2.4: Viele tragbare Sensoren benutzen einen künstlichen, farblich homogenen Hintergrund, der hinter den Trauben positioniert wird. Dies ermöglicht eine einfache Unterscheidung von Traube und Hintergrund und vermeidet, dass neben der Traube weitere Bestandteile der Weinrebe erfasst werden, aus Aquino et al. [2017], modifiziert.

halb-automatisch. Diese Methoden verletzen daher die Kriterien der Automation und der Datenerhebung im Hochdurchsatz und sind daher für die Belange in NoViSys ebenfalls nicht geeignet.

Soll die Erfassung des Traubenbestandes hingegen nicht-invasiv im Hochdurchsatz erfolgen, muss die Datenerhebung zwingend **indiskriminativ** geschehen. Indiskriminativ bedeutet in diesem Kontext, dass keine vorherige Selektion der vom Sensor zu erfassenden Objekte durch einen Anwender stattfindet. Stattdessen umfassen die Daten neben den Trauben alle weiteren erfassten Objekte im Sichtfeld des Sensors. Dazu gehören u. a. Blätter, Äste und sonstige Bestandteile des Drahtgerüsts, sowie Objekte des Hintergrundes wie dem Himmel oder zurückliegende Anbaureihen.

Nur Methoden, welche auf einer indiskriminativen Datenerhebung basieren, erschließen erst die Möglichkeit, Daten aus der Bewegung und automatisiert mittels einer mobilen Sensorplattform zu erheben. Im Folgenden seien daher ausschließlich jene Methoden hinsichtlich der Erfüllung der Kriterien aus Kapitel 1 eingeordnet,

welche Daten indiskriminativ und unter Feldbedingungen erheben (Tabelle 2.1, Feld rechts unten). Die dort aufgelisteten Arbeiten erzeugen in ihren Einsatzbereichen gute Ergebnisse. Sie erfüllen die in Kapitel 1 genannten Kriterien für die Vorhaben in NoViSys jedoch nur in Teilen:

Nicht-Invasivität und Nahbereich: Alle vorgestellten Arbeiten verwenden nicht-invasive, optische Sensoren für den Einsatz im Nahbereich.

Hochaufgelöst: Alle Arbeiten nutzen Sensoren, deren Auflösungen potentiell die Zählung einzelner Beeren ermöglichen. In einigen Arbeiten werden die Daten ausschließlich klassifiziert [Fernández et al., 2013; Klodt et al., 2015], teilweise werden ausschließlich Trauben zur Schätzung des Ertrages genutzt [Dey et al., 2012; Diago et al., 2012; Liu and Whitty, 2015; Luo et al., 2016; Marinello et al., 2016; Santos et al., 2017]. Die Vernachlässigung der Beerenzählung ist jedoch nicht optimal, da Beeren als kleinster Gewichtsanteil der Traube in Summe das Gesamtgewicht der Traube bilden und erst die Beerenanzahl eine präzise Ertragsschätzung ermöglicht.

Maßstab: Zur Einführung eines metrischen Maßstabes werden teilweise bekannte Maßstäbe im Sichtfeld des Sensors positioniert, anschließend automatisch erkannt und ein Maßstabsfaktor für die Überführung in metrische Einheiten berechnet [Herrero-Huerta et al., 2015; Kicherer et al., 2015a; Santos et al., 2017]. Diese Vorgehensweise ist jedoch für die automatische Datenerhebung im Hochdurchsatz ungeeignet, da Maßstäbe in regelmäßigen Abständen durch Personal in den Anbaureihen platziert werden müssen. Eine automatische Möglichkeit ist die Nutzung eines Stereo-Kamerasystems mit bekannter Basislinie oder die Verwendung eines Laserscanners [Marinello et al., 2016; Mirbod et al., 2016].

Hochdurchsatz und Automation: Alle Arbeiten nutzen Sensoren, die theoretisch eine Datenerhebung aus der Bewegung heraus ermöglichen. Voll funktionsfähige und automatisierte Sensorplattformen für die Datenaufnahme im Hochdurchsatz unter Feldbedingungen sind aufgrund der vielfältigen Herausforderungen bislang jedoch kaum realisiert worden. Vollständige Sensorplattformen wurden von Kicherer et al. [2015a], Aquino et al. [2018] und Nuske et al. [2014b] entwickelt. Diese Sensorplattformen sind in der Lage, die Daten aus der Bewegung heraus zu erheben und erreichen teilweise Geschwindigkeiten bis zu 7 km/h (Aquino). In den restlichen Publikationen werden stattdessen mittels tragbarer Sensoren, Stativen oder provisorischen Transportmitteln vollständige Sensorplattformen und die Ausrichtung der Sensoren zur Anbaureihe lediglich simuliert und Konzeptionsvorschläge für die Realisierung einer Sensorplattform unterbreitet. Die Datenaufnahme geschieht dadurch aus dem Stand heraus durch einen Anwender. Diese Arbeiten können somit keine Aussagen über die Schwierigkeiten bei der Datenerhebung mit einer tatsächlichen Sensorplattform wie bspw. Platzproblemen, durch Berührung induzierte Bewegung der Weinreben oder die Geschwindigkeit der Datenerhebung treffen.

Im Gegensatz zur Betrachtung isolierter Trauben gewinnt bei der indiskriminativen Datenerhebung die automatische Klassifikation der Daten als elementarer Automatisierungsschritt an Bedeutung, da die Daten neben Trauben/Beeren weitere Objekte wie Blätter, Äste, Himmel oder zurückliegende Anbaureihen enthalten. Alle Arbeiten nutzen dazu Klassifikationsalgorithmen basierend auf Farb-, Geometrie- oder Texturdeskriptoren, mit Ausnahme von Herrero-Huerta et al. [2015], welche Trauben manuell aus der Punktwolke segmentieren müssen. Die anschließende Vermessung und Zählung der Trauben und Beeren anhand der klassifizierten Daten geschieht hingegen für alle Arbeiten automatisch.

Standardisierte Aufnahmebedingungen: Unter Feldbedingungen können wechselnde Beleuchtungsverhältnisse oder Regen die Daten stark beeinflussen. Objekte des Hintergrunds wie der Himmel oder zurückliegende Anbaureihen erschweren zusätzlich die Klassifikation der Daten. Wechselnde Beleuchtungsverhältnisse werden von Luo et al. [2016], Liu and Whitty [2015] und Diago et al. [2012] durch entsprechend angepasste Klassifikationsmodelle berücksichtigt.

In keinem Ansatz werden der Sensor und die betrachteten Weinreben vor äußeren Verhältnissen abgeschirmt. Künstliche Beleuchtungsquellen zur Standardisierung der Beleuchtungsverhältnisse werden zwar von Nuske et al. [2014b], Mirbod et al. [2016], Aquino et al. [2018], Fernández et al. [2013] und Kicherer et al. [2015a]

verwendet, ein interner Kamerablitze von Liu and Whitty [2015]. Nuske et al. [2014b], Mirbod et al. [2016], Aquino et al. [2018] und Kicherer et al. [2015a] erreichen standardisierte Aufnahmebedingungen durch die Datenerhebung bei Nacht, wodurch eine stets identische Ausleuchtung, sowie ein durchgehend schwarzer Hintergrund erreicht wird, der sich von Weinreben leicht unterscheiden lässt. Diese Vorgehensweise schränkt jedoch die Einsatzmöglichkeiten dieser Methoden ein und schirmt nicht vor Regen ab.

Vollständigkeit: Um eine hohe Geschwindigkeit der Datenerhebung zu erreichen, müssen die Anbaureihen indiskriminativ erfasst werden. Dabei ist zu berücksichtigen, dass Trauben durch Blätter, Äste und andere Trauben aus der Perspektive des Sensor mitunter verdeckt sind, da sie nicht händisch isoliert werden können. Abbildung 2.5 demonstriert das Szenario in einer typischen Spalieranlage. Trauben können dicht beieinander



Abbildung 2.5: Ein typisches Szenario in einer Spalieranlage. Trauben sind durch andere Traube, Blätter und Äste verdeckt. Dicht liegende Trauben erschweren ihre Unterscheidung. Ausleuchtung, Schattenwurf und Regen erzeugen variierende Erscheinungsbilder der Weinreben.

liegen, sich gegenseitig verdecken oder bei ausgebildetem Blattwerk durch Blätter und Äste verdeckt werden. Diese Verdeckungen erschweren die vollständige Erfassung des Traubenbestandes und erhöhen die Unsicherheit der Ertragsschätzung. Verdeckungen müssen daher durch mehrere Perspektiven entlang der Horizontalen und Vertikalen minimiert werden, um möglichst viele Trauben erfassen zu können. Mit zunehmenden Perspektiven des Sensors auf die Anbaureihe verringert sich der Verdeckungsgrad und der Traubenbestand kann vollständiger erfasst werden.

Mehrere Perspektiven erzeugen dabei überlappende Bildinhalte und bedingen daher zwangsläufig eine räumliche Registrierung der Daten, um Mehrfachzählungen in den redundanten Bildinhalten zu vermeiden. Aquino et al. [2018], Nuske et al. [2014b], Herrero-Huerta et al. [2015], Mirbod et al. [2016] Santos et al. [2017] und Dey et al. [2012] nehmen sequentielle Bildfolgen aus unterschiedlichen Perspektiven mit stark überlappenden Bildinhalten auf. Mehrfachzählungen werden in diesen Methoden entweder über 3D-Rekonstruktion durch Multi-View-Stereo Methoden [Dey et al., 2012; Herrero-Huerta et al., 2015; Santos et al., 2017], Zusammenfügung von drei Bildern zu einem einzigen Bild [Aquino et al., 2018] oder über die Betrachtung aller Bilder, die einen definierten Abschnitt der Anbaureihe abbilden [Nuske et al., 2014b]. Bei Nuske et al. [2014b] wird ausschließlich das Bild mit der höchsten detektierten Beerenanzahl zur Zählung genutzt. Die Ergebnisse der übrigen Bilder werden verworfen, weshalb die erhobenen Daten einen vergleichsweise höheren Verdeckungsgrad besitzen. Die übrigen oben genannten Methoden vermeiden Mehrfachzählungen, die Kamera verbleibt während der Datenerhebung aber in nur einer Höhe, außer bei Herrero-Huerta et al. [2015], wo die sequentielle Bildaufnahme mit zwei weiteren Höhen wiederholt, um zusätzliche Perspektiven auf die Anbaureihen einzunehmen und den Einfluss von Verdeckungen weiter zu minimieren.

Alle anderen Methoden basieren auf einer einzigen Perspektive für einen Abschnitt der Anbaureihe und sind

daher stärker von Verdeckungen beeinflusst.

Mit Ausnahme von Liu and Whitty [2015] wird in allen Arbeiten außerdem stets nur eine Seite der Anbaureihe erfasst.

Zeitnahe Referenzwerte: In den Arbeiten von Nuske et al. [2014b], Mirbod et al. [2016] und Aquino et al. [2018] werden zeitnah erhobene Referenzwerte verwendet, um für die vom Sensor nicht erfassten Trauben und Beeren für den gesamten betrachteten Abschnitt der Anbaureihe für die Ertragsschätzung zu berücksichtigen. In den Arbeiten von Herrero-Huerta et al. [2015], Diago et al. [2012] und Santos et al. [2017] werden ebenfalls zeitnah erhobene Referenzwerte genutzt. Die geschieht jedoch nur, um die Zahl der bereits detektierten Beeren bzw. Trauben zu korrigieren. Vom Sensor nicht erfasste Trauben werden nicht berücksichtigt, wodurch die Anzahl der Ertragsparameter unvollständig bleibt.

Lokalisierung: Georeferenzierung wurde bisher in Nuske et al. [2014b], Mirbod et al. [2016], Aquino et al. [2018] und Kicherer et al. [2015a] mittels GPS erreicht. Die automatische Herstellung eines räumlichen Bezugs der Ergebnisse wurde in den übrigen Arbeiten bislang nicht berücksichtigt.

2.3.2 Vorgehensweise zur Lösung der drei Teilaufgaben

Unter Bezug auf die Defizite der oben besprochenen Arbeiten hinsichtlich der in Kapitel 1 formulierten Kriterien wird zur Lösung der gestellten Aufgabe dieser Arbeit folgendermaßen vorgegangen:

1. **Datenerhebung:** In dieser Arbeit werden zwei mobile Sensorplattformen, der *PHENObot* und der *Phenoliner*, für den Einsatz unter Feldbedingungen und die Datenerhebung im Hochdurchsatz untersucht. Der PHENObot wurde ursprünglich im Zuge der Projekte PHENOVines und CROPSense [PHENOVines, 2017; Project CROP.SENSE, 2017] entwickelt. Der Phenoliner ist eine prototypische mobile Sensorplattform, die im Rahmen des NoViSys-Projektes entstand und auf dessen Belange ausgerichtet ist.

Zur Datenerhebung werden RGB-Kameras verwendet. Kameras sind kleine, flexible, kostengünstige Sensoren und erreichen eine genügend hohe Objektauflösung für die Erfassung der kleinmaßstäbigen Beeren. Die RGB-Kameras nehmen aus der Fahrt heraus RGB-Bilder der Anbaureihen auf, die später zur Rekonstruktion farbiger 3D-Punktwolken der Anbaureihen mittels einer Structure-from-Motion/Multi-View-Stereo Methode (SfM/MVS) [Furukawa et al., 2015; Hartley and Zisserman, 2004] genutzt werden. Die RGB-Kameras nehmen dabei eine Vielzahl an Perspektiven entlang der Fahrtrichtung und der Vertikalen ein. Dadurch werden Verdeckungen der Trauben durch eine vollständigere Betrachtung der Weinreihen kompensiert und gleichzeitig die 3D-Geometrie der Trauben und Beeren erhalten.

Durch Einsatz von GPS erhalten die Punktwolken unmittelbar einen metrischen Maßstab. Strategien zur Schaffung standardisierter Aufnahmebedingungen wie die Abschirmung der Kameras und Weinreben vor äußeren Einflüssen, die Verwendung künstlicher Beleuchtung sowie eines künstlichen Hintergrundes werden untersucht. Die gewonnenen Punktwolken werden zur Schätzung der Ertragsparameter und des Ertrages genutzt.

2. **Datenklassifikation:** In den 3D-Punktwolken werden mittels überwachter Klassifikation automatisch diejenigen 3D-Punkte detektiert, die Trauben abbilden. Dabei werden sowohl die Farbe als auch die Geometrie der Weinreben ausgenutzt. Mehrere geometrische, wie farbliche Deskriptoren, sowie Kombinationen beider Arten von Deskriptoren werden hinsichtlich ihrer Eignung zur Klassifikation der Weinreben in Voruntersuchungen evaluiert und der geeignetste Deskriptor identifiziert.

Für den identifizierten Deskriptor werden Untersuchungen zur Optimierung seiner Unterscheidungskraft durchgeführt. Dabei wird eine Graph-Cut Methode verwendet, welche räumliche Nachbarschaften zwischen den bereits klassifizierten 3D-Punkten ausnutzt, um die Klassifikation zu verbessern.

3. **Quantifikation:** Um die Ertragsparameter und den Ertrag zu bestimmen, werden diejenige klassifizierten 3D-Punkte genutzt, welche Trauben abbilden. Zunächst werden die für die Kameras sichtbaren Beeren und Trauben detektiert. Sichtbare Beeren und Trauben werden über ein 3D-Connected Components Verfahren in Kombination mit einem eigenen Algorithmus detektiert, der auf Kugelapproximationen basiert. Die zusätzlich dritte Dimension der Punktwolke erleichtert 1. die Unterscheidung dicht benachbarter Trauben, die sich im 2D-Raum räumlich häufig überlappen und ermöglicht die Schätzung der Beerengröße über den Kugelradius. Um von der Anzahl sichtbarer Beeren bzw. Trauben auf die Ertragsparameter (Gesamtanzahl aller Beeren bzw. zu Trauben) zu schließen, werden Kalibrierwerte aus Referenzwerten berechnet und mit der Anzahl sichtbarer Beeren und Trauben verrechnet. Basierend auf den geschätzten Ertragsparametern wird der Ertrag geschätzt.

3. Methodische Grundlagen

In dieser Arbeit wird das Verständnis zweier grundlegender Methoden benötigt:

1. 3D-Rekonstruktion von Objekten mittels 2D-Bildern der RGB-Kameras des PHENObots und des Phenoliners über *Structure-from-Motion und Multi-View-Stereo Methoden*
2. *Überwachte Klassifikation* der Punktwolken anhand von *Deskriptoren*

Aufgrund ihrer methodischen Bedeutung und des nötigen Umfangs der Erläuterungen werden sie in diesem eigenen Kapitel behandelt¹.

3.1 3D-Rekonstruktion der Anbaureihen mittels einer SfM/MVS-Methode

Die kommerzielle Software *Pix4DMapper* (Pix4D SA, 1015 Lausanne, Switzerland) wird zur Rekonstruktion farbiger, georeferenzierter Punktwolken aus den Bildern der RGB-Kameras auf den Sensorplattformen verwendet. Sie wurde bereits in Rose et al. [2015] für die 3D-Phänotypisierung von Pflanzen erfolgreich genutzt. Ihr Algorithmus basiert auf zwei sequentiellen Schritten: 1) *Structure-from-Motion* (SfM)- und 2) Multi-View-Stereo (MVS) [Küng et al., 2012; Remondino et al., 2014; Strecha et al., 2012]. Zur 3D-Rekonstruktion werden Bilder aus unterschiedlichen Perspektiven mit hoher inhaltlicher Bildüberlappung von mind. 60% benötigt [Tsingas, 1992]. Eine hohe Überlappung der Bilder ist für Pix4DMapper essentiell, um identische Bildpunkte in den Bildern identifizieren zu können [Furukawa et al., 2015]. Aus mehreren Bildperspektiven können Traubenbestände erfasst werden, die von anderen Perspektiven ansonsten verdeckt wären, wodurch der Traubenbestand vollständiger erfasst wird.

Im SfM-Schritt werden die Bilder zueinander räumlich registriert und anschließend im MVS-Schritt zur 3D-Rekonstruktion der Bildinhalte verwendet. Bei der Überführung vom 2D-Raum der Bilder in den 3D-Raum der Punktwolke werden dadurch redundante Erfassungen des gleichen Traubenbestandes verhindert. SfM/MVS ist daher besonders für die Erfassung der komplexen Morphologie der Weinreben geeignet. Detaillierte Studien und Erläuterungen zu weiteren SfM/MVS Lösungen neben Pix4DMapper sind bspw. in Hartley and Zisserman [2004], Remondino et al. [2014] oder Knapitsch et al. [2017] zu finden.

Im Folgenden wird die 3D-Rekonstruktion mittels SfM/MVS beschrieben. Dabei wird auf ein exzessives Formelwerk verzichtet. Stattdessen wird dem Leser die grundlegende Funktionsweise erläutert, die es ihm ermöglichen soll, potentielle Fehlerquellen und Bewertungsmöglichkeiten mittels SfM/MVS rekonstruierter Punktwolken zu verstehen und den späteren Äußerungen zu folgen.

3.1.1 Bildpunkte und innere Kameraparameter

Ein Bild besteht aus Bildpunkten, die kartesische 2D-Koordinaten im *Bildkoordinatensystem* besitzen. Das Bildkoordinatensystem spannt sich entlang der zueinander orthogonalen Bildachsen der Länge u bzw. v in Pixeleinheiten $[pel]$ auf und besitzt im obersten linken Pixel des Bildes seinen Ursprung. Die Bildpunktkoordinaten eines Bildpunktes $\mathbf{x}'_n$ seien über $\mathbf{x}'_n = [x', y']_n$, $n = 1, \dots, (u \cdot v)$ definiert. Die wichtigsten sog. *innere Kameraparameter* der verwendeten Kamera sind:

1. Kamerakonstante F : Orthogonaler Abstand des Projektionszentrums O von der Sensorebene der Kamera. Sie ist ein berechneter Wert, der in etwa der Brennweite der Kameralinse entspricht.

¹Weitere im Zuge dieser Arbeit verwendete Methoden werden in darauffolgenden Kapiteln an der Stelle ihrer ersten Anwendung erläutert, um Funktionsweise und Anwendung für den Leser didaktisch nah zusammen zu halten.

2. Hauptpunkt $[x'_0, y'_0]$: Projektion des Projektionszentrums in die Sensorebene der Kamera. Dieser liegt etwa in der Sensormitte.

Abbildung 3.1 stellt die Lage der inneren Kameraparameter relativ zum Bildkoordinatensystem dar.

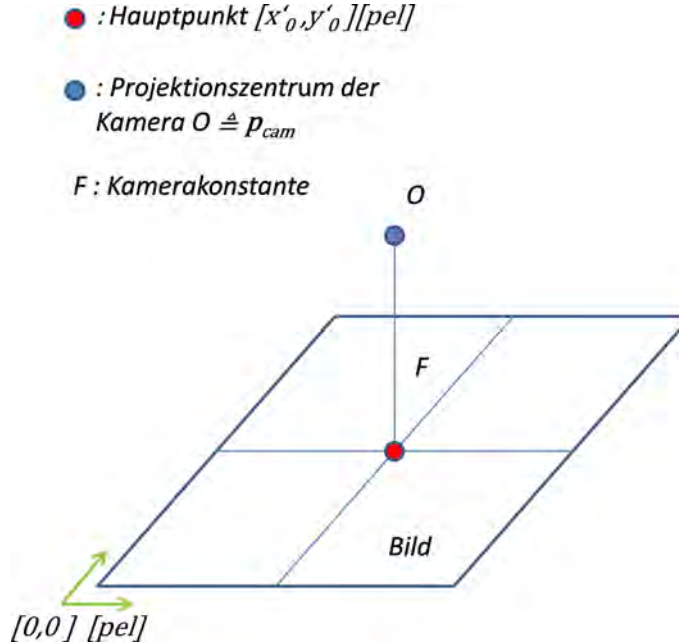


Abbildung 3.1: Lage des Hauptpunktes und der Kamerakonstanten in Relation zum Ursprung des Bildkoordinatensystems (grüne Pfeile).

3.1.2 Structure-from-Motion anhand homologer Bildpunkte

Structure-from-Motion (SfM) bezeichnet die Bestimmung von 3D-Koordinaten im Objektraum aus den 2D-Koordinaten sog. *homologer Bildpunkte* im Bildkoordinatensystem. Als homologe Bildpunkte bezeichnet man Bildpunkte, die den gleichen Objektpunkt in mehreren Bildern unterschiedlicher Position und Perspektive abbilden. Als Objektpunkt bezeichnet man den 3D-Punkt der physischen, vom Pixel abgebildeten Stelle der betrachteten Szene. Eine Voraussetzung für die Identifikation einer ausreichenden Anzahl homologer Bildpunkte ist dabei eine horizontale und vertikale inhaltliche Bildüberlappung, sowie eine heterogene Textur der Bilder zur Wiedererkennung und Unterscheidung von homologen Bildpunkten. Diese lassen sich automatisch detektieren, bspw. durch Nutzung des *SIFT-Algorithmus* [Lowe, 2004]. Der gesamte 2D-Raum des Bildes wird dabei nach besonders distinktiven Bildpunkten, sog. *Keypoints* [Lowe, 2004], durchsucht, die über Ähnlichkeitsmaße in anderen Bildern wieder erkannt werden können.

Zunächst werden die relativen Orientierungen $\mathbf{R}_{cam_j}$ und die Positionen der Kameras, sowie die 3D-Koordinaten der identifizierten homologen Bildpunkte (Keypoints) im Objektraum über Näherungswerte der inneren Kameraparameter und Kamerapositionen berechnet. Die benötigten Näherungswerte der inneren Kameraparameter ergeben sich aus der Herstellerangabe der Brennweite und den Bildkoordinaten der Mitte der Sensorebene. Für die Rekonstruktion ist es daher nötig, entweder stets dieselbe Kamera oder identische Kameras mit identischen inneren Kameraparametern zu verwenden. Abbildung 3.2 illustriert eine Weinrebe und einen sukzessiven Bildverband mit J-Bildern mit homologen Bildpunkten. Die Bilder des PHENObots und des Phenoliners (s. Kapitel 4 und 5) werden über GPS georeferenziert. Die GPS-Positionen der Kameras $\mathbf{p}_{cam_j} = [X, Y, Z]_{cam_j} \in \mathbb{R}_3$, $j = 1, \dots, J$ mit J-Bildern gelten als Näherungswerte der Kamerapositionen. Kameras werden paarweise anhand homologer Bildpunkte relativ zueinander orientiert [Hartley and Zisserman, 2004]. Die Bestimmung der Näherungswerte der 3D-Koordinate eines von beiden Kameras

beobachteten, homologen Bildpunktes erfolgt über Triangulation der sog. *Projektionsstrahlen*. Ein Projektionsstrahl wird durch den Bildpunkt $\mathbf{x}_{nj}$ eines Objektpunktes und das Projektionszentrum O_j des jeweiligen Bildes j festgelegt. Seine räumliche Lage wird über die relative Orientierung und Position der Kamera bestimmt (Abbildung 3.3). Die paarweise relative Orientierung der Kameras und Triangulation der homologen Bildpunkte geschieht für alle Kameras.

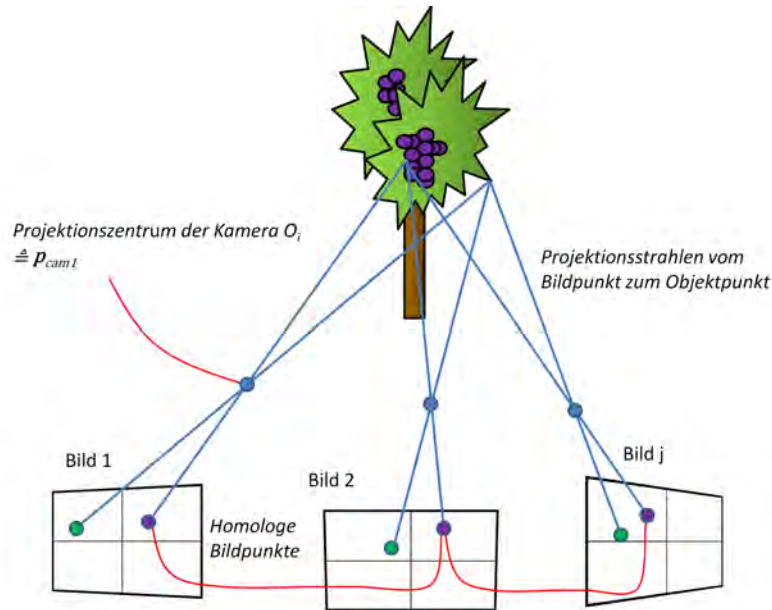


Abbildung 3.2: Ein sequentieller Bildverband mit homologen Bildpunkten (lila: ein Objektpunkt auf einer Traube; grün: ein Objektpunkt auf einem Blatt), den Kamerapositionen $\mathbf{p}_{cam_j}$ der Bilder und den Projektionsstrahlen zum Objektpunkt, die durch das Projektionszentrum O_j der jeweiligen Kamera laufen.

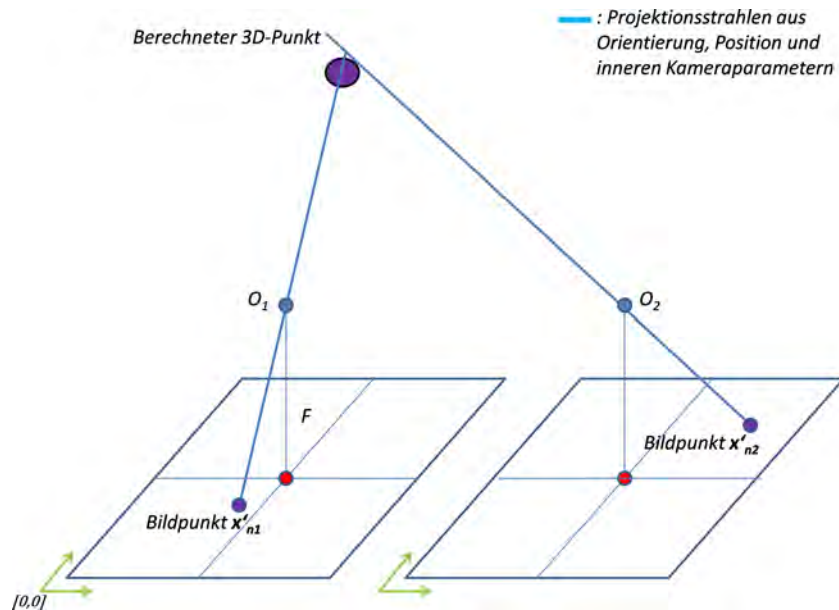


Abbildung 3.3: Die Projektionsstrahlen zum Objektpunkt verlaufen vom Bildpunkt $\mathbf{x}'_{nj}$ durch das Projektionszentrum O_j und werden durch Orientierung und Kameraposition räumlich festgelegt. Die Koordinaten des Objektpunktes werden durch Triangulation der Projektionsstrahlen aller Bildpunkte bestimmt, die ihn abbilden.

In einer *Bündelblockausgleichung* [Hartley and Zisserman, 2004] werden anschließend die Orientierung $\mathbf{R}_{cam_j}$ und Position $\mathbf{p}_{cam_j}$ der Kamera zu allen Aufnahmezeitpunkten, die inneren Kameraparameter, sowie die 3D-Koordinaten der homologen Bildpunkte nach der Methoden der kleinsten Quadrate ausgeglichen. Die ausgeglichene Orientierung $\mathbf{R}_{cam_j}$ und Position $\mathbf{p}_{cam_j}$ einer Kamera ist synonym mit Orientierung und Position des Projektionszentrums O_j der Kameralinse.

Durch Bestimmung der Näherungswerte der Kameraposition über GPS wird unmittelbar ein metrischer Maßstab für die Punktwolke festgelegt und die Punktwolke im übergeordneten Koordinatensystem des GPS-Systems gelagert. Eine Maßstabsfestlegung aus bekannten Objekten in der Punktwolke wie in Herrero-Huerta et al. [2015] ist nicht erforderlich. Die 3D-Punkte der homologen Bildpunkte werden als *dünne Punktwolke* bezeichnet, da sie ausschließlich aus den Objektpunkten der homologen Bildpunkte besteht.

3.1.3 Multi-View-Stereo: Verdichtung der dünnen Punktwolke

Die aus SfM resultierende dünne Punktwolke besteht nur aus wenigen Punkten mit geringem geometrischen Detailgrad, da nicht jeder einzelne Bildpunkt $\mathbf{x}'_{nj}$ des Bildes j , sondern lediglich die 3D-Koordinaten der Objektpunkte der homologen Bildpunkte berechnet werden. Nach der Bündelblockausgleichung sind die geschätzte Orientierung $\mathbf{R}_{cam_j}$, Kameraposition jedes Bildes im GPS-System $\mathbf{p}_{cam_j}$ und innere Kameraparameter gegeben. Multi-View-Stereo Techniken rekonstruieren mit Hilfe dieser Informationen eine *dichte Punktwolke* und nutzen dazu alle Bildpunkte $\mathbf{x}'_{nj}$ eines Bildes. Dazu werden wie bei SfM homologe Bildpunkte benötigt.

MVS-Techniken nutzen das Prinzip der *Epipolar Geometrie*, um diese homologen Bildpunkte in mehreren Bildern zu suchen. Abbildung 3.4 stellt die Epipolar Geometrie anhand eines Beispiels mit zwei Bildern dar. Die Projektionszentren O_1 bzw. O_2 , sowie der gesuchte Objektpunkt $\mathbf{p}$ bilden die Eckpunkte der sog. *Epipolarebene*. Die Raumlinie zwischen den Projektionszentren O_1 und O_2 wird als Basis bezeichnet. Der Epipol ist der Punkt, wo die Basis die jeweilige Bildebene durchstößt. Der Bildpunkt $\mathbf{x}'_1$ im linken Bild sei bekannt. Dann bildet die Projektion des Projektionsstrahls $\mathbf{x}'_1 - O_1$ auf die rechte Bildebene die sog. *Epipolarlinie* (grüne Linie). Der ebenfalls mit dem Objektpunkt $\mathbf{p}$ korrespondierende, gesuchte Bildpunkt $\mathbf{x}'_2$ des rechten Bildes muss sich auf dieser Epipolarlinie befinden. Die Suche nach homologen Bildpunkten, welche bei SfM-Techniken im 2D-Raum des Bildes stattfindet, reduziert sich daher bei MVS auf die 1D-Suche auf der Epipolarlinie der benachbarten Bilder. Die Suche auf der Epipolarlinie geschieht dabei über Ähnlichkeitsmaße der Intensitäten der Bildpunkte [Furukawa et al., 2015].

Nach Identifikation der homologen Bildpunkte wird die Position des Objektpunktes über Triangulation ihrer Projektionsstrahlen [Furukawa et al., 2015] berechnet. Zur Triangulation müssen mindestens drei homologe Bildpunkte gefunden worden sein. Der Schnittwinkel α der Projektionsstrahlen, sowie die Häufigkeit der Beobachtungen eines Objektpunktes beeinflussen die Genauigkeit seiner Positionsbestimmung. Je vielfältiger die Schnittwinkel sind, desto sicherer kann seine 3D-Position trianguliert werden. Hingegen führen bei hauptsächlich spitzen oder stumpfen Schnittwinkeln kleine Ungenauigkeiten bei der Berechnung der Projektionsstrahlen bereits zu Verfälschungen der Positionskoordinaten [Witte and Schmidt, 2006].

Innere Kameraparameter, Orientierung und Position der Bilder müssen zudem möglichst genau bestimmt sein, um die geometrischen Bedingungen der Epipolar Geometrie zu erfüllen. Projektionsstrahlen homologer Bildpunkte schneiden sich andernfalls nicht ausreichend genau in einem Punkt und die Genauigkeit der Punktbestimmung sinkt.

Heterogene Bildinhalte mit unterschiedlichen Intensitäten sind für die Unterscheidung von Bildpunkten bei der Suche auf der Epipolarebene vorteilhaft, homogene Bildinhalte erschweren hingegen die korrekte Identifikation. Die Rekonstruktion scheitert daher vor allem bei gleichfarbigen Oberflächen, bei denen Bildpunkte als homolog identifiziert werden können, die nicht tatsächlich miteinander korrespondieren. Geschieht dies, wird ein 3D-Punkt erzeugt, der keinem tatsächlichen Objektpunkt entspricht. Sind die gleichfarbigen Bildinhalte großflächig, können auf diese Weise Punktregionen mit vielen Punkten entstehen, die keine Entsprechung in der Realität besitzen.

Dynamische Bildinhalte verfälschen ebenfalls die Rekonstruktion. Zu starke Bewegungen der Objekte während der Aufnahme führen dazu, dass mehrfache Instanzen des gleichen Objektpunktes an jeweils unterschiedlichen Positionen erzeugt werden.

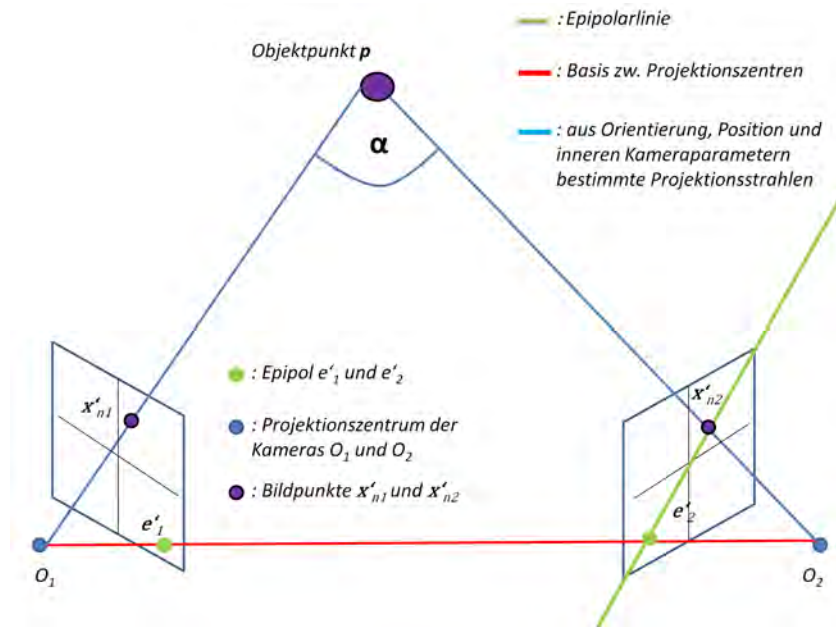


Abbildung 3.4: Prinzip der Epipolar Geometrie anhand zweier Bilder, in denen der gleiche Objektpunkt p beobachtet wurde.

Die triangulierten 3D-Koordinaten der beobachteten Objektpunkte werden als letztes nach der Methode der kleinsten Quadrate ausgeglichen. Die so entstehende dichte Punktwolke sei mit $\mathcal{P}$ bezeichnet.

3.1.4 Bewertung der Punktwolke anhand des Rückprojektionsfehlers

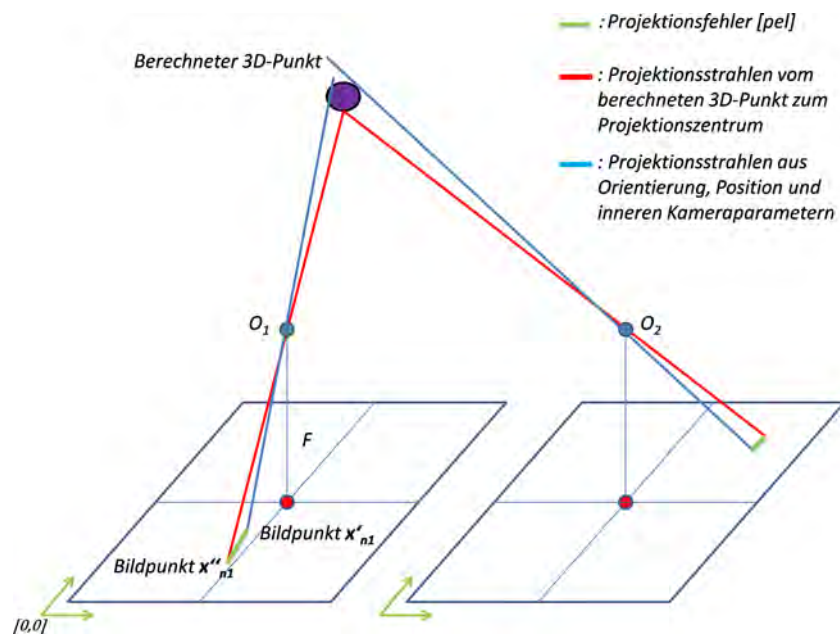


Abbildung 3.5: Bestimmung des Rückprojektionsfehlers durch Rücktransformation des Objektpunktes in die Bildkoordinatensysteme.

Der durchschnittliche sog. *Rückprojektionsfehler* gibt die Übereinstimmung von beobachtetem Bildpunkt und rekonstruiertem 3D-Punkt an. Er wird bei der Ausgleichung minimiert und ist ein Indikator für die Genauigkeit der inneren Geometrie der Punktwolke [Hartley and Zisserman, 2004]. Pix4DMapper berechnet den durchschnittlichen Rückprojektionsfehler d aus allen generierten 3D-Punkten in $\mathcal{P}$. Er wird in Pixeleinheiten angegeben und sollte $< 1[pe]$ sein [Furukawa et al., 2015]. Abbildung 3.5 illustriert das Prinzip des Rückprojektionsfehlers anhand zweier Bilder.

Jeder 3D-Punkt $\mathbf{p}_i \in \mathcal{P}$ der Punktwolke wurde initial als Bildpunkt an einer 2D-Position $\mathbf{x}'_{nj}$ auf der Sensorebene eines Bildes j abgebildet. Es gilt $\mathbf{p} \hat{=} \mathbf{x}'_{nj}$. Mittels der Orientierung und der Position der Bilder im übergeordneten Koordinatensystem wird der Projektionsstrahl zwischen $\mathbf{p}_i$ und O_j berechnet und vom 3D-Punkt zurück durch das Projektionszentrum O_j des Bildes an die Position $\mathbf{x}''_{nj}$ im Bildkoordinatensystem projiziert. Der Rückprojektionsfehler $d(\mathbf{x}', \mathbf{x}'')_{nj}$ eines 3D-Punktes ergibt sich über die euklidische Distanz im Bildkoordinatensystem: $d(\mathbf{x}', \mathbf{x}'')_{nj} = \|\mathbf{x}' - \mathbf{x}''\|_{nj}$. Aus allen individuellen Rückprojektionsfehlern jedes 3D-Punktes in jedes Bild j , in dem er beobachtet wurde, wird ein Durchschnittswert des Rückprojektionsfehlers für die gesamte Punktwolke berechnet.

3.1.5 Zusammenfassung

Die nach SfM/MVS entstandene Punktwolke wird mit $\mathcal{P} = [\mathbf{p}_1, \dots, \mathbf{p}_N]$ bezeichnet, mit N -Punkten $\mathbf{p}_i = [X, Y, Z]_i \in \mathbb{R}_3$, $i = 1, \dots, N$. Jedem Punkt in $\mathcal{P}$ sind die vom Bildsensor erfassten Rot-Grün-Blau-Werte (RGB) des abgebildeten Objekts zugewiesen. Des Weiteren werden die ausgeglichenen Kamerapositionen $\mathbf{p}_{cam_j} = [X, Y, Z]_{cam_j} \in \mathbb{R}_3$, $j = 1, \dots, J$ mit J -Bildern in UTM-Koordinaten erhalten. Die Punktwolke $\mathcal{P}$ wird nach Abschluss der Rekonstruktion in einem *kd-tree* organisiert, um mit hoher Geschwindigkeit nach nächsten Nachbarn der 3D-Punkte suchen zu können [Bentley, 1990]. Homogen gefärbte Objekte oder sich bei der Bildaufnahme stark bewegende Objekte verhindern die korrekte Zuweisung homologer Bildpunkte, wodurch die Rekonstruktion korrumpiert werden kann. Die rekonstruierten Objekte sind des Weiteren mit zufälligen Abweichungen behaftet, die für eine korrekte Abbildung der Objektgeometrie reduziert werden müssen.

3.2 Datenklassifikation

3.2.1 Grundlegende Klassifikationsstrategie

Die Punktwolke $\mathcal{P}$ enthält neben den Trauben weitere Objekte. Diese müssen von den Trauben unterschieden werden. Eine manuelle Trennung von Trauben aus der Punktwolke ist langwierig, erfordert die volle Aufmerksamkeit des Anwenders und ist damit kostspielig. Alle Punkte $\mathbf{p}_i \in \mathcal{P}$, die Trauben abbilden, sollen daher automatisiert klassifiziert werden, um für die Schätzung der Ertragsparameter genutzt werden zu können.

Klassifikation bezeichnet im Allgemeinen die auf distinktiven Merkmalen basierte Einordnung von Objekten in semantische Kategorien, im Weiteren als *Klassen* bezeichnet. Der Mensch erfasst Klassen primär durch Vorwissen und seiner Fähigkeit, Objekte anhand seiner Sinneseindrücke zu kategorisieren. Da in NoViSys ausschließlich die Trauben von Interesse sind, werden die Klassen **Traube** und **Rest** definiert, aus denen sich die Gesamtheit aller Objekte in der Punktwolke konstituiert. Die Klasse **Rest** beinhaltet dabei alle Bestandteile der Anbaureihen, die keine Trauben sind. Dazu zählen u. a. Holz, Blätter, Draht, Stützpfiler oder Markierungsmaterial.

Ein *automatischer Klassifikator* benötigt für die Unterscheidung semantischer Klassen quantifizierbare Merkmale. Ein 3D-Punkt besitzt zunächst lediglich seine XYZ-Koordinaten im kartesischen Koordinatensystem als Merkmale. Für die Klassifikation sind die Punktkoordinaten als alleinige Klassenmerkmale für die Zugehörigkeit zur Klasse **Traube** jedoch nicht ausreichend distinktiv.

Für die Klassifikation wird jeder Punkt der Punktwolke daher mit einem sog. *Deskriptorvektor* ausgestattet. Ein Deskriptorvektor enthält $1 \times f$ sog. *Deskriptoreinträge*, wobei die Dimension von f durch die Methode seiner Berechnung bestimmt wird. Ein Deskriptorvektor enthält quantifizierte Merkmale eines Punktes, anhand derer er von Punkten mit anderen Deskriptoreinträgen unterschieden werden kann. Um Verwechslungen vorzubeugen, wird in dieser Arbeit von *Deskriptor* gesprochen, wenn die *Methode der Berechnung* gemeint ist und von *Deskriptorvektor*, wenn der *berechneten Vektor* gemeint ist.

Im Verlauf dieser Arbeit werden sowohl *geometrische* wie *farbliche* Deskriptoren, sowie Kombinationen beider Arten von Deskriptoren hinsichtlich ihrer Eignung für die Klassifikation von Weinreben untersucht. Geometrische Deskriptoren quantifizieren die Geometrie innerhalb eines definierten räumlichen Umkreises, farbliche Deskriptoren die Farbe von Objekten.

3D-Punkte innerhalb einer ähnlichen geometrischen Umgebung wie bspw. auf der Wölbung einer Beere werden ähnliche geometrische Deskriptorvektoren besitzen, die sich aber distinktiv von 3D-Punkten auf der lokalen Ebene eines Blattes oder auf den zylinderförmigen Ästen unterscheiden. Jeder Punkt der Punktwolke besitzt zudem durch die farbigen Bilder RGB-Merkmale des abgebildeten Objektpunktes. 3D-Punkte, die Trauben abbilden, besitzen ähnliche Farben, unterscheiden sich aber farblich von Ästen, Blättern und anderen Objekten der Anbaureihe. Die zwei definierten Klassen besitzen damit distinktive Merkmale, die sich anhand dieser Deskriptorvektoren beschreiben lassen.

Geometrie- und farbbasierte Deskriptoren haben unterschiedliche Vor- und Nachteile. Geometriebasierte Deskriptoren sind abhängig von der geometrischen Exaktheit der 3D-Oberflächen und sensibel gegenüber zufälligen Abweichungen der 3D-Punkte von der Soll-Geometrie. Des Weiteren können Geometrien bspw. durch überlappende Objekte oder durch eine zu großen Entfernung zur Kamera geometrisch korrumpiert sein. Geometrische Deskriptoren sind jedoch unabhängig von der Farbe der Objekte. Farbbasierte Deskriptoren hingegen berechnen sich unabhängig von der Objektgeometrie. Sie variieren mit wechselnden Lichtverhältnissen oder dem aktuellen Farbausdruck des Objekts. Insbesondere Pflanzen sind hochgradig variabel und ändern ihren Farbausdruck je nach Reifegrad, Gesundheitsstatus oder durch Einsatz von Fungiziden [Diago et al., 2012]. Die Kombination von sowohl geometrie-als auch farbbasierten Deskriptoren erscheint daher sinnvoll, um einen robusteren Deskriptor zu generieren, in dem die jeweiligen Vor- und Nachteile ausgeglichen werden.

In dieser Arbeit werden insgesamt sechs Deskriptoren untersucht: die geometrischen Deskriptoren *Surface-Feature-Histograms (SFH)* und *Saliency-Features (SF)*, der farbliche HSV-Deskriptor, der aus den RGB-Werten der 3D-Punkte bestimmt wird (s. Kapitel 4.1.4), sowie die jeweilige Kombination der SFH- und SF-Deskriptoren mit dem HSV-Deskriptor durch Konkatenation zum SFHc- bzw. SFc-Deskriptor. Ein weiterer untersuchter Deskriptor ist *CSHOT (Color-Signature-of-Histograms-of-Orientations)*, der den geometrischen Deskriptor *SHOT* mit einem farblichen Deskriptor im *CIELab-Farbraum* kombiniert. Diese Deskriptoren gehören zu den populärsten Deskriptoren und wurden bereits für die Klassifikation von Punktwolken urbaner Szenen, Pflanzen und Weinreben eingesetzt (s. jeweiliges Kapitel zum Deskriptor).

In den Untersuchungen mit dem PHENObot (Kapitel 4) wird sich zeigen, dass mit der Kombination von SFH- und HSV-Deskriptor zum SFHc-Deskriptor die besten Klassifikationsergebnisse erreicht werden. Der SFHc-Deskriptor wird daher als Deskriptor für den Schritt der Datenklassifikation in den Praxisuntersuchungen zur Schätzung der Ertragsparameter und des Ertrages genutzt. Die Berechnung des SFH-Deskriptors ist komplex und wird im Folgenden daher detailliert beschrieben. Die Berechnung der HSV-, SF(c), und CSHOT-Deskriptoren ist zum weiteren Verständnis kurz umrissen, für ausführliche Erläuterungen sei auf die zitierte Literatur verwiesen.

Der überwachte, probabilistische Klassifikator *Incremental Import Vector Machine (IVM)* [Roscher et al., 2012a] wird für die Klassifikation der Punktwolken verwendet. Die IVM berechnet ein sog. *Klassifikationsmodell*, welches aus *Trainingsdaten*, die Objekte beider Klassen repräsentieren, und den Deskriptorvektoren der Trainingsdaten generiert wird. Mittels dieses Klassifikationsmodells werden für alle Punkte der zu klassifizierenden Punktwolke anhand ihrer Deskriptorvektoren die wahrscheinlichen Klassenzugehörigkeiten für einer der beiden Klassen **Trauben** oder **Rest** bestimmt. Dieses erste Klassifikationsergebnis wird als *initiale Klassifikation* bezeichnet.

Für die initiale Klassifikation wurden die Punkte unabhängig voneinander betrachtet, d. h. ohne Berücksichtigung der Klassenzugehörigkeit benachbarter 3D-Punkte eines Punktes. Die Graph-Cut-Software [Boykov et al., 2001] *GCO* [Delong et al., 2012; Kolmogorov and Zabini, 2004] wird verwendet, um die räumlichen Nachbarschaften der Punkte und ihre wahrscheinlichen Klassenzugehörigkeiten aus der initialen Klassifikation zur Verbesserung der Klassifikationsergebnisse auszunutzen und eine homogene Klassifikation zu erreichen. Der Vorgang wird als *Klassenglättung* bezeichnet.

3.2.2 Erläuterung der untersuchten Deskriptoren

Im Folgenden werden die geometrischen und farblichen Deskriptoren erläutert. Zunächst wird die Berechnung des farblichen HSV-Deskriptors aus RGB-Werten geschildert. Darauf folgt die Berechnung der geometrischen Deskriptoren. Sämtliche hier beschriebenen geometrischen Deskriptoren sind lokale Deskriptoren, welche die Geometrie innerhalb eines nutzerdefinierten Bereichs um einen Quellpunkt beschreiben. Die Berechnung der geometrischen Deskriptoren basiert auf der Berechnung von Normalenvektoren für alle Punkte $\mathbf{p}_i \in \mathcal{P}$. Die Vorgehensweise für die Berechnung von Normalenvektoren wird daher als Erstes erläutert. Im Anschluss erfolgt die Erläuterung der SF-, SFc-, CSHOT, SFH und SFHc-Deskriptoren.

3.2.2.1 Hue-Saturation-Value

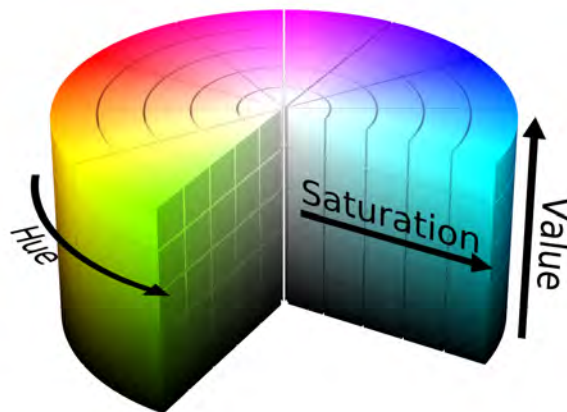


Abbildung 3.6: Der HSV-Raum repräsentiert Farbe über zylindrische Koordinaten für die Hue-Saturation-Value-Werte. Bildautor: SharkD.²

Im Bereich der Phänotypisierung von Pflanzen und der Bildverarbeitung wird häufig der *HSV-Raum* anstelle des RGB-Raums verwendet [Bora et al., 2015; Dhondt et al., 2013; Font et al., 2015]. Bei der Transformation von RGB-Werten in den HSV-Raum werden Farben über zylindrische Koordinaten basierend auf den drei Koordinatenwerten Hue (Farbe), Saturation (Farbsättigung) und Value (Helligkeit) repräsentiert.

Abbildung 3.6 stellt das zylindrische Koordinatensystem des HSV-Raums dar. Die HSV-Werte besitzen einen Wertebereich von $[0..1]$. Die Hue-Koordinate entspricht dem Winkel auf dem Farbkreis, die Saturation-Koordinate einem Anteil des Radius des Zylinders und die Value-Koordinate einem Anteil der Höhe des Zylinders. Die HSV-Werte werden als farblicher Deskriptor $\mathbf{hsv} = [\text{Hue Saturation Value}]$ der 3D-Punkte verwendet.

RGB-Werte $[r, g, b]$ eines Punktes werden über ihre Minima- und Maxima-Werte und durch Fallunterscheidungen in HSV-Werte transformiert:

²CC BY-SA 3.0 Lizenz: <https://creativecommons.org/licenses/by-sa/3.0/deed.en>

Algorithmus 3.1 : Pseudocode der HSV Transformation.**Data** : RGB Werte $[r, g, b]$ eines 3D-Punktes**Result** : HSV Werte $[H, S, V]$ des 3D-Punktes

```

1 Bestimme Maxima  $max_{rgb}$  und Minima  $min_{rgb}$  im Wertebereich von  $[r, g, b]$ ;
2 Value =  $max_{rgb}$ ;
3  $\Delta = max_{rgb} - min_{rgb}$ ;
4 if  $max_{rgb} \neq 0$  then
5   Saturation =  $\Delta / max_{rgb}$ 
6 end
7 ;
8 if  $max_{rgb} == 0$  then
9    $r = g = b == 0$ ;
10  dann folgt: Saturation = 0, Hue = -1, Value ist undefiniert.
11 end
12 ;
13 if  $max_{rgb} \in Rot$  then
14   Hue =  $60 \cdot ((g - b) / \Delta)$ 
15 end
16 ;
17 if  $max_{rgb} \in Grün$  then
18   Hue =  $60 \cdot (2 + (b - r) / \Delta)$ 
19 end
20 ;
21 if  $max_{rgb} \in Blau$  then
22   Hue =  $60 \cdot (4 + (r - g) / \Delta)$ 
23 end
24 ;
25 if Hue < 0 then
26   Hue = Hue + 360°.
27 end
28 ;

```

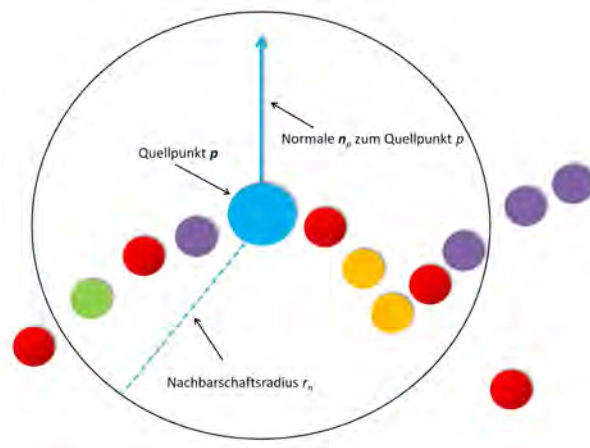
3.2.2.2 Berechnung von Normalenvektoren

Abbildung 3.7: Der Normalenvektor n_p zum Punkt p (maßstäblich hervorgehoben), der sich aus der PCA seiner Nachbarnpunkte im Radius r_n ergibt.

Die Normalenvektoren aller Punkte in $\mathbf{p}_i \in \mathcal{P}$ werden durch *Principal Component Analysis* (PCA) berechnet. Seien ein Quellpunkt $\mathbf{p} \in \mathcal{P}$ und die Menge von K -Nachbarpunkten $\mathcal{P}^k = \mathbf{q}_1 \dots \mathbf{q}_i, i = 1 - K$ im Nachbarschaftsradius r_n um $\mathbf{p}$ gegeben.

Die Suche nach den Nachbarpunkten geschieht über den kd-tree, in dem $\mathcal{P}$ organisiert ist. Es gilt, dass $\mathcal{P}^k = \{q_i | \|\mathbf{q}_i - \mathbf{p}\| \leq r_n\}$. Abbildung 3.7 illustriert die Approximation des Normalenvektors eines Quellpunktes $\mathbf{p}$ durch Betrachtung seiner lokalen Nachbarschaft. Die Wahl des Radius r_n ist ein gemeinsamer Berechnungsparameter aller hier verwendeten geometrischen Deskriptoren und hat große Bedeutung für die Ausprägung der Deskriptoreinträge. Im Kapitel 4.2.1 wird näher darauf eingegangen.

Der Schwerpunkt der Nachbarpunkte $\mathcal{P}^k$ wird über $\bar{\mathbf{q}} = \frac{1}{K} \sum_{i=1}^K \mathbf{q}_i$ berechnet. Anschließend folgt die Berechnung der Kovarianzmatrix $\mathbf{C} \in \mathbb{R}^{3 \times 3}$ der Nachbarpunkte $\mathcal{P}^k$ um $\mathbf{p}$:

$$\mathbf{C} = \frac{1}{K} \sum_{i=1}^K (\mathbf{q}_i - \bar{\mathbf{q}})^T (\mathbf{q}_i - \bar{\mathbf{q}}). \quad (3.1)$$

Die Kovarianzmatrix ist symmetrisch und positiv-semi-definit, ihre Eigenwerte $\lambda_3 \geq \lambda_2 \geq \lambda_1 \geq 0$ sind reelwertige Zahlen $\in \mathcal{R}$. Die zugehörigen Eigenvektoren formen einen orthogonalen Rahmen. Der Eigenvektor $\mathbf{v}_1$ zum kleinsten Eigenwert λ_1 ist eine *Approximation des Normalenvektors* $\mathbf{n}_p$ zum Quellpunkt $\mathbf{p}$. Das Vorzeichen, d. h. Orientierung, von $\mathbf{n}_p$ ist dabei zunächst zufällig positiv oder negativ. Ist jedoch eine Beobachtungsposition $\mathbf{v}_p \hat{=} \mathbf{p}_{cam}$ gegeben, von dem aus die Oberfläche um den Punkt $\mathbf{p}$ erfasst wurde, lässt sich die Orientierung des Normalenvektors zu $\mathbf{v}_p$ ausrichten, wenn die folgende Bedingung geprüft wird:

$$\mathbf{n}_p \cdot (\mathbf{v}_p - \mathbf{p}) > 0 \quad (3.2)$$

Ist die Bedingung nicht erfüllt, wird das Vorzeichen des Normalenvektors umgekehrt und damit auf die Beobachtungsposition ausgerichtet. Die geschilderte Vorgehensweise zur Berechnung und Orientierung der Normalenvektoren aller N -Punkte der Punktwolke $\mathcal{P}$ wird für jeden Punkt $\mathbf{p}_i \in \mathcal{P}$ wiederholt, der dabei jeweils als Quellpunkt $\mathbf{p}$ fungiert.

3.2.2.3 Saliency Features

Saliency Features (SF) wurden erstmalig von Lalonde et al. [2006] für die Klassifikation natürlicher Objekte im Kontext der autonomen Navigation von terrestrischen Robotern vorgestellt. Dey et al. [2012] nutzten sie bereits für die Klassifikation von farbigen Punktwolken von Weinreben in die Klassen **Äste**, **Blätter** und **Früchte** und sollen daher für die Vorhaben in NoViSys untersucht werden.

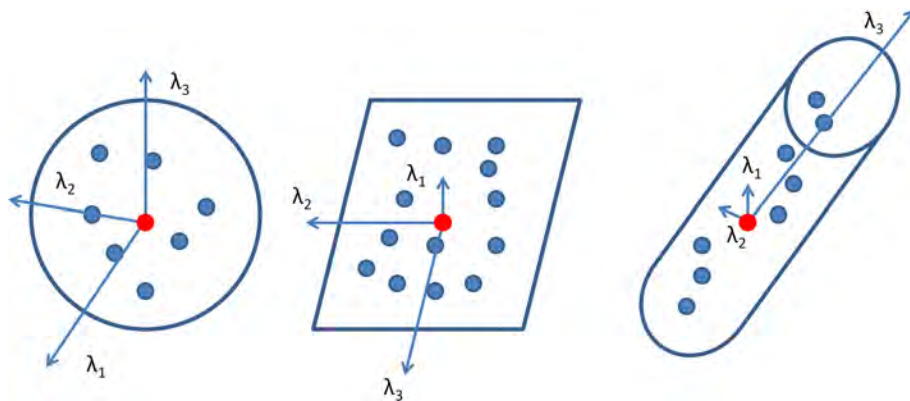


Abbildung 3.8: Die drei geometrischen Erscheinungsformen Volumen (links), Ebenheit (Mitte) und Linearität (rechts). Die Größenordnung der Eigenwerte $\lambda_3 - \lambda_1$ zu einem Quellpunkt (rot) innerhalb einer Punktregion $\mathcal{P}^k$ (blau) wird durch die Länge des entsprechenden Eigenvektors angezeigt.

SF-Deskriptoren beschreiben die drei geometrische Erscheinungsformen *Ebenheit*, *Linearität* und *Volumen* einer betrachteten Punktreion $\mathcal{P}^k$ im Radius r_n um einen Quellpunkt $\mathbf{p}$ anhand der Eigenwerte $\lambda_3, \lambda_2, \lambda_1 \geq 0$, die sich aus der PCA der betrachteten Punktreion ergeben (s. Kapitel 3.2.2.2). Die dominante Anordnung der 3D-Punkte in $\mathcal{P}^k$ bestimmt dabei die Größenordnung der Eigenwerte (Abbildung 3.8).

Die drei geometrischen Erscheinungsformen berechnen sich über:

$$\text{Volumen} = \lambda_3, \text{ wobei } \lambda_3 \approx \lambda_2 \approx \lambda_1 \quad (3.3)$$

$$\text{Ebenheit} = \lambda_2 - \lambda_1, \text{ wobei } \lambda_3 \approx \lambda_2 \gg \lambda_1 \quad (3.4)$$

$$\text{Linearität} = \lambda_3 - \lambda_2, \text{ wobei } \lambda_3 \gg \lambda_2 \approx \lambda_1. \quad (3.5)$$

Ein SF-Deskriptor für einen 3D-Punkt im Nachbarschaftsradius r_n besitzt demnach die Dimension $f = 3$. Für Trauben wird als voluminöse Geometrie ein dominanter Wert für Volumen erwartet, für die ebenen Blätter und die geradlinige Geometrie der Äste hohe Werte für Ebenheit bzw. Linearität.

Es wird den Ausführungen von Dey et al. [2012] gefolgt und für jeden 3D-Punkt $\mathbf{p}_i \in \mathcal{P}$ die SF-Deskriptoren $\mathbf{sf}_1$ und $\mathbf{sf}_2$ mit zwei verschiedenen Werten für den Nachbarschaftsradius r_n berechnet:

$$\mathbf{sf}_1 = [\text{Volumen Ebenheit Linearität}] \text{ mit Nachbarschaftsradius } r_n \quad (3.6)$$

$$\mathbf{sf}_2 = [\text{Volumen Ebenheit Linearität}] \text{ mit Nachbarschaftsradius } (r_n \times 2.) \quad (3.7)$$

Die einzelnen Deskriptorvektoren $\mathbf{sf}_1$ und $\mathbf{sf}_2$ werden anschließend zu einem einzigen Deskriptorvektor $\mathbf{sf}$ zusammen gefasst:

$$\underset{(1 \times 6)}{\mathbf{sf}} = \left[\underset{(1 \times 3)}{\mathbf{sf}_1} \quad \underset{(1 \times 3)}{\mathbf{sf}_2} \right]. \quad (3.8)$$

Das geschilderte Vorgehen wird für jeden 3D-Punkt $\mathbf{p}_i \in \mathcal{P}$ ausgeführt. Der Wertebereich wird für jeden Deskriptoreintrag auf $[0..1]$ normiert.

3.2.2.4 Saliency-Features-Color

Die SF- und HSV-Deskriptoren werden zum SFC-Deskriptor mit der Dimension $f = 6 + 3 = 9$ konkateniert:

$$\underset{(1 \times 9)}{\mathbf{sfc}} = \left[\underset{(1 \times 6)}{\mathbf{sf}} \quad \underset{(1 \times 3)}{\mathbf{hsv}} \right]. \quad (3.9)$$

Der Wertebereich aller Deskriptoreinträge des SFC-Deskriptorvektors ist auf $[0..1]$ normiert. Da sowohl farb- als auch geometriebasierte Deskriptoren kombiniert werden, die sich auf unterschiedliche Weise berechnen, ist die Normierung für die spätere Klassifikation von Bedeutung, um maßstäblich größeren Werten keine zu hohe Gewichtung einzuräumen [Kumar, 2017; Patki and Kelkar, 2013].

3.2.2.5 Color-Signature-of-Histograms-of-Orientations

Der Deskriptor CSHOT (Color-Signature-of-Histograms-of-Orientations) [Tombari et al., 2011] basiert auf dem geometrischen Deskriptor SHOT (Signature-of-Histograms-of-Orientations) [Tombari et al., 2010]. SHOT wurde ursprünglich als rein geometrischer Deskriptor für die Objekterkennung entwickelt [Alhamzi et al., 2015], wurde jedoch auch bereits in Behley et al. [2012] für die Klassifikation von Punktwolken urbaner Objekte im Kontext der Roboter Navigation eingesetzt. CSHOT wurde als Erweiterung von SHOT um einen farbbasierten Deskriptor im CIELab-Farbraum [Commission Internationale De l'Eclairage, 2017] speziell für die Objekterkennung bei farbigen Punktwolken entwickelt und ist daher ein geeigneter Kandidat für die Klassifikation der farbigen Punktwolken von Weinreben. Es erfolgt zunächst die Erläuterung von SHOT und anschließend die Erweiterung von SHOT um den Deskriptor im CIELab-Farbraum zum CSHOT-Deskriptor.

Für die Berechnung des SHOT-Deskriptors wird zunächst eine PCA für alle 3D-Punkte in $\mathcal{P}$ durchgeführt. Für einen betrachteten Quellpunkt $\mathbf{p}$ und seinen Nachbarnpunkten $\mathcal{P}^k$ im Nachbarschaftsradius r_n wird aus den Eigenvektoren der gewichteten Kovarianzmatrix ein lokales Koordinatensystems mit $\mathbf{p}$ als Ursprung berechnet [Behley et al., 2012; Tombari et al., 2010]. Das Koordinatensystem wird anschließend im sog. *Histogrammradius* r_H um $\mathbf{p}$ in ein rotationsinvariantes, kugelförmiges Raster um $\mathbf{p}$ unterteilt. Dabei muss stets $r_H \geq r_n$ gelten. Der Histogrammradius r_H ist damit neben r_n der zweite Berechnungsparameter von CSHOT.

Die Menge der im Raster enthaltenen Nachbarnpunkte von $\mathbf{p}$ seien mit $\mathcal{P}^h$ bezeichnet. Das Raster besitzt acht Unterteilungen entlang des Azimuts, zwei Unterteilungen entlang der Elevation und zwei Unterteilungen entlang des Histogrammradius r_H . Insgesamt werden dadurch 32 einzelne Sektoren $S_i, i = 1..32$ erhalten. Abbildung 3.9 skizziert den Vorgang der Raumunterteilung. Folgend wird für jeden Punkt $\mathbf{p}_i \in \mathcal{P}$ einzeln vorgegangen.

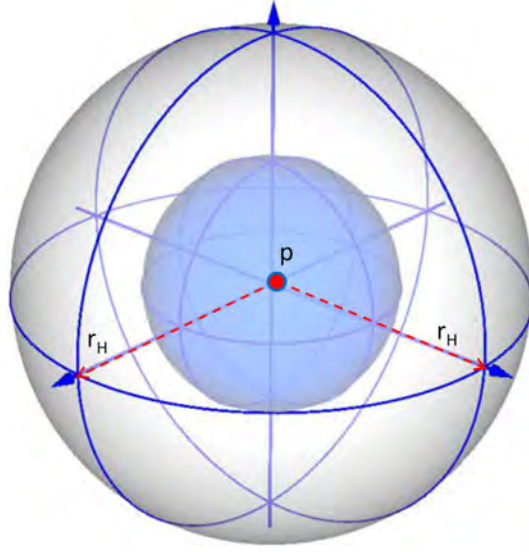


Abbildung 3.9: Skizze der Unterteilung des Raumes in ein kugelförmiges Raster um $\mathbf{p}$ (roter Punkt). Es werden nur vier azimutale Unterteilungen dargestellt. Der Radius r_H ist rot gestrichelt dargestellt, aus Tombari et al. [2010], modifiziert

Für jeden Sektor des Rasters wird der Winkel α zwischen der Normalen $\mathbf{n}_{\mathbf{p}}$ des Quellpunktes $\mathbf{p}$ und den Normalen aller Nachbarnpunkte in $\mathcal{P}^h$ im jeweiligen Sektor S_i bestimmt und ein Histogramm der α -Werte mit $b = 11$ -Intervallen berechnet. Die einzelnen Histogramme jedes Sektors werden zuletzt zum SHOT-Deskriptor der Dimension $f = 32 \times 11 = 352$ zusammengefügt. Die Anzahl der Intervalle b ist der dritte Berechnungsparameter von CSHOT.

Zur Erweiterung von SHOT zu CSHOT werden die RGB-Werte aller 3D-Punkte in $\mathcal{P}$ zunächst in den CIELab-Farbraum transformiert. In diesem wird Farbe anhand der drei Werte für *Lightness*, $\mathbf{a}$ für die Grün-Rot Komponente und $\mathbf{b}$ für die Blau-Gelb-Komponente der Farbe quantifiziert. Man erhält somit für jeden Punkt $\mathbf{p}$ einen Vektor $\mathbf{l}_p = [L \ a \ b]$ im CIELab-Farbraum. Es werden erneut alle Sektoren einzeln betrachtet. Für jede Kombination des Quellpunktes $\mathbf{p}$ und eines Nachbarnpunktes $\mathbf{q}$ im betrachteten Sektor S_i wird die L1-Norm $L(\mathbf{l}_p, \mathbf{l}_q)$ berechnet. Für alle Werte von $L(\mathbf{l}_p, \mathbf{l}_q)$ im jeweiligen Sektor wird anschließend ein Histogramm mit $b = 31$ -Intervallen berechnet. Die Histogramme der einzelnen Sektoren S_i werden anschließend zu einem Deskriptorvektor **ciel** der Dimension $f = 32 \times 31 = 992$ konkateniert.

CSHOT ergibt sich durch Zusammenfügen des SHOT- und CIEL-Deskriptorvektors:

$$\underset{(1 \times 1344)}{\mathbf{cshot}} = \begin{bmatrix} \underset{(1 \times 352)}{\mathbf{shot}} & \underset{(1 \times 992)}{\mathbf{ciel}} \end{bmatrix}. \quad (3.10)$$

In dieser Arbeit wird die CSHOT-Implementation der Point Cloud Library [Willow Garage, 2017] genutzt.

3.2.2.6 Surface Feature Histograms

Die Surface Feature Histograms (SFH) basieren auf den Fast-Point-Feature-Histograms (FPFH) von Rusu [2009], der sie zur Klassifikation künstlicher 3D-Objekte für die Navigation von Robotern in Innenräumen oder für die Registrierung von Punktwolken einsetzte. Paulus et al. [2013] adaptierte sie für die Phänotypisierung von komplexen 3D-Objekten wie Pflanzen und konnte erstmalig erfolgreich ihre Eignung für die Klassifikation von Weinreben, Gerste oder Weizenähren demonstrieren [Paulus et al., 2014a].

Die SFH beschreiben die Geometrie innerhalb einer gewählten Nachbarschaft $\mathcal{P}^k \in \mathcal{P}$ um einen Quellpunkt $\mathbf{p} \in \mathcal{P}$ durch Berechnung von Winkeln zwischen den Normalenvektoren der betrachteten Punkten. Wie CSHOT sind sie von der Wahl der Berechnungsparameter r_n , r_H und b abhängig. Zur Berechnung der SFH werden neben der Berechnung von Normalenvektoren zwei Schritte durchlaufen:

1. Berechnung der sog. *Simplified-Point-Feature-Histograms* (SPFH) aus Winkelkombinationen der Normalenvektoren aller Punkte.
2. Generierung der SFH durch gewichtete Mittelung der SPFH.

Berechnung der SPFHs

Im ersten Schritt werden die Simplified-Point-Feature-Histograms (SPFH) für jeden 3D-Punkt in $\mathbf{p}_i \in \mathcal{P}$ berechnet, der jeweils als Quellpunkt $\mathbf{p}$ fungiert. Es wird dazu ein Quellpunkt $\mathbf{p}$, die Menge seiner H-Nachbarnpunkte $\mathbf{q}_i \in \mathcal{P}^h, i = 1..H$ im Histogrammradius r_H sowie ihre Normalenvektoren herangezogen, wobei stets $r_H \geq r_n$ gelten muss. Es gilt, dass $\mathcal{P}^h = \{\mathbf{q}_i \mid \|\mathbf{q}_i - \mathbf{p}\| \leq r_H\}$. Abbildung 3.10 illustriert den Nachbarschaftsradius r_H und die Paarbildung von Quellpunkt und Nachbarnpunkten.

Für jedes Punktepaar $(\mathbf{p}, \mathbf{q}_i \in \mathcal{P}^h)$ werden anschließend drei Merkmale f_1, f_2, f_3 berechnet. Diese sind ein Maß der Verkippen α, ϕ, θ des Normalenvektors $\mathbf{n}_{q_i}$ gegenüber dem sog. *Darboux-Rahmen*. Abbildung 3.11 illustriert das geometrische Verhältnis der Normalenvektoren von Quell- und Nachbarnpunkt, den Darboux-Rahmen und die Lage der Winkel α, ϕ und θ .

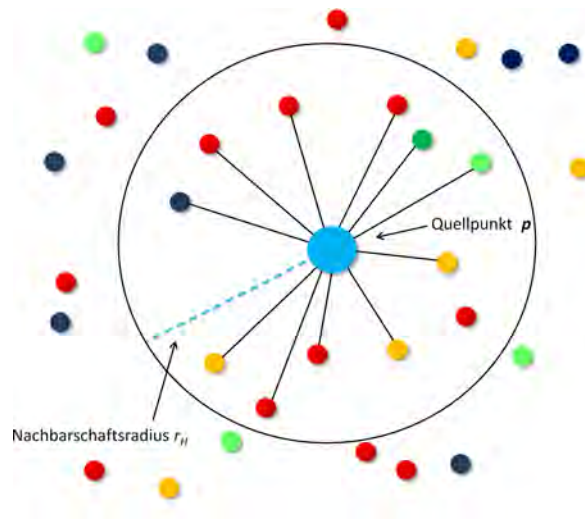


Abbildung 3.10: Die zur Generierung des SPFH betrachteten Punkte der Nachbarschaft $\mathcal{P}^h$ um $\mathbf{p}$. Jede Kombination zwischen $\mathbf{p}$ und den Nachbarn $\mathbf{q}_i \in \mathcal{P}^h$, angedeutet durch die Verbindungsgeraden, wird zur Berechnung dreier Merkmale aus den zugehörigen Normalenvektoren herangezogen.

Der Darboux-Rahmen $\mathbf{u}-\mathbf{v}-\mathbf{w}$ fungiert als rechtwinkliges, lokales Referenzsystem für jedes Punktpaar $(\mathbf{p}, \mathbf{q}_i)$. Er ergibt sich über

$$\mathbf{u} = \mathbf{n}_p \quad (3.11)$$

$$\mathbf{v} = \frac{(\mathbf{q}_i - \mathbf{p}) \times \mathbf{u}}{\|(\mathbf{q}_i - \mathbf{p}) \times \mathbf{u}\|} \quad (3.12)$$

$$\mathbf{w} = \mathbf{u} \times \mathbf{v}, \quad (3.13)$$

wobei $\|\dots\|$ die euklidische Länge des Vektors indiziert.

Der Darboux-Rahmen wird lokal für jede Kombination $(\mathbf{p}, \mathbf{q}_i)$ berechnet und ist daher invariant gegenüber Rotation und Position der Punktwolke $\mathcal{P}$ [Rusu, 2009]. Da sich die Merkmale auf den Darboux-Rahmen beziehen, sind sie ebenfalls invariant gegenüber Rotation und Position der Punktwolke.

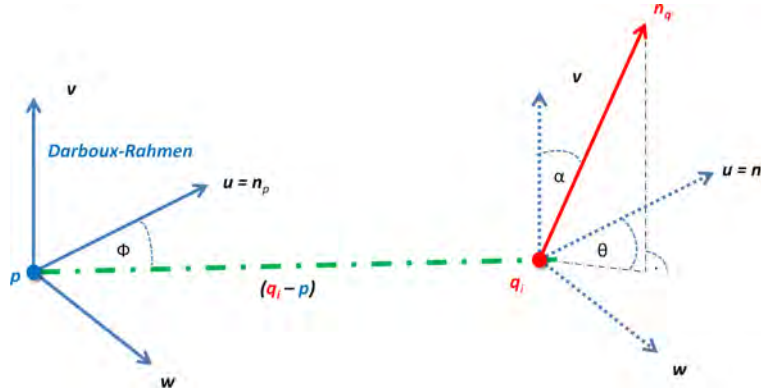


Abbildung 3.11: Der Darboux-Frame (blaue Pfeile) zum Quellpunkt $\mathbf{p}$ (blau) und dessen Nachbarpunkt $\mathbf{q}_i$ (rot). Die Winkel α und θ ergeben sich zwischen dem Darboux-Rahmen in $\mathbf{q}_i$ (blau gestrichelt) und dem Normalenvektor $\mathbf{n}_{q_i}$ des Nachbarpunktes (roter Pfeil).

Die Merkmale f_1, f_2, f_3 werden unter Bezug auf den Darboux-Rahmen nach folgenden Formeln berechnet:

$$f_1 = \mathbf{v} \cdot \mathbf{n}_{q_i} \quad (3.14)$$

$$f_2 = \frac{(\mathbf{u} \cdot (\mathbf{q}_i - \mathbf{p}))}{\|(\mathbf{q}_i - \mathbf{p})\|} \quad (3.15)$$

$$f_3 = \text{atan}(\mathbf{w} \cdot \mathbf{n}_{q_i}, \mathbf{u} \cdot \mathbf{n}_{q_i}) \quad (3.16)$$

Der Winkel α ist der Winkel zwischen der Normalen des Nachbarpunktes $\mathbf{n}_{q_i}$ und $\mathbf{v}$, ϕ ist der Winkel zwischen der Normalen des Quellpunktes $\mathbf{n}_p$ und der Verbindungsgeraden zwischen $\mathbf{p}$ und $\mathbf{q}_i$. Die sie beschreibenden Merkmale f_1 und f_2 werden über das Skalarprodukt zweier normalisierter Vektoren beschrieben. Sie sind daher tatsächlich die Kosinus-Werte der Winkel zwischen jeweils beiden verwendeten Vektoren. Ihr Wertebereich liegt somit bei $[-1..0..1]$. Das Merkmal f_3 stellt direkt den Winkel θ zwischen der Projektion von $\mathbf{n}_{q_i}$ auf die durch $\mathbf{n}_p$ und $\mathbf{w}$ aufgespannte Ebene dar. Sein Wertebereich reicht somit von $[-\frac{\pi}{2}..0..\frac{\pi}{2}]$.

Die Wertebereiche der Merkmale $f_1 - f_3$ werden anschließend in jeweils b -Intervalle unterteilt. Für jedes Punktpaar $(\mathbf{p}, \mathbf{q}_i)$ wird geprüft, in welches Intervall die Werte der Merkmale $f_1 - f_3$ jeweils fallen. Bei drei Merkmalen $f_1 - f_3$ sind b^3 Intervallkombinationen möglich. Für diese Arbeit werden $b = 5$ Intervalle verwendet, was $f = 125$ Kombinationsmöglichkeiten entspricht. Diese Anzahl hat sich bereits bei Paulus et al. [2014a] als geeignet für die Klassifikation von Weinreben erwiesen³.

³Diese Vorgehensweise unterscheidet sich von dem Vorgehen bei Berechnung der FPFHs in Rusu [2009], wo für jedes Merkmal $f_1 - f_3$ ein eigenes Histogramm mit $b = 11$ Intervallen gebildet und diese drei Histogramme zu einem Deskriptor mit der Dimension $f = 3 \cdot 11 = 33$ konkateniert werden.

Die jeweils belegten Intervalle werden für alle Punktpaare $(\mathbf{p}, \mathbf{q}_i \in \mathcal{P}^h)$ bestimmt. Die Häufigkeit jeder Intervallkombination innerhalb von $\mathcal{P}^h$ wird anschließend gezählt und in einem Histogramm mit $b^3 = 125$ Bins aufsummiert. Dabei entspricht jede Intervallkombination einem Bin des Histogramms. Eine gegebene Intervallkombination erhöht das entsprechende Bin um den Betrag $\frac{1}{H}$, wodurch der Wertebereich des Histogramms zwischen $[0..1]$ liegt. Die Höhe jedes Bins hängt somit lediglich von ihrem relativen Anteil an der Gesamtanzahl der insgesamt vorhandenen Intervallkombinationen ab. Das Histogramm ist daher invariant gegenüber der Anzahl der betrachteten Punkte in $\mathcal{P}^h$.

Dieses Histogramm ist das Surface-Point-Feature-Histogram mit der Dimension $f = 125$. Die SPFHs aller Punkte in $\mathcal{P}$ werden auf die oben geschilderte Weise nacheinander bestimmt.

Berechnung der SFHs

Im zweiten Schritt werden die SFHs durch gewichtete Mittelung der SPFHs berechnet. Im Radius r_H um einen Quellpunkt $\mathbf{p}$ werden dazu alle SPFHs der Nachbarpunkte in $\mathcal{P}^h$ betrachtet. In Abhängigkeit ihrer euklidischen Distanz d zu $\mathbf{p}$ erhalten die SPFHs der Nachbarpunkte folgende Gewichtung:

$$wb_1(k) = 1 - (0.5 + \frac{d}{r_H} \cdot 0.5) \quad (3.17)$$

Der Quellpunkt $\mathbf{p}$ erhält komplementär die von wb_1 abhängige Gewichtung

$$wb_2(k) = (1 - wb_1(k)) \quad (3.18)$$

Der jeweilige Deskriptorvektor **sfh** ergibt sich über

$$\underset{(1 \times 125)}{\mathbf{sfh}}(p) = \mathbf{spfh}(p) + \frac{1}{k} \sum_{i=1}^k wb_1 \cdot \mathbf{spfh}(q_i) + wb_2 \cdot \mathbf{spfh}(p). \quad (3.19)$$

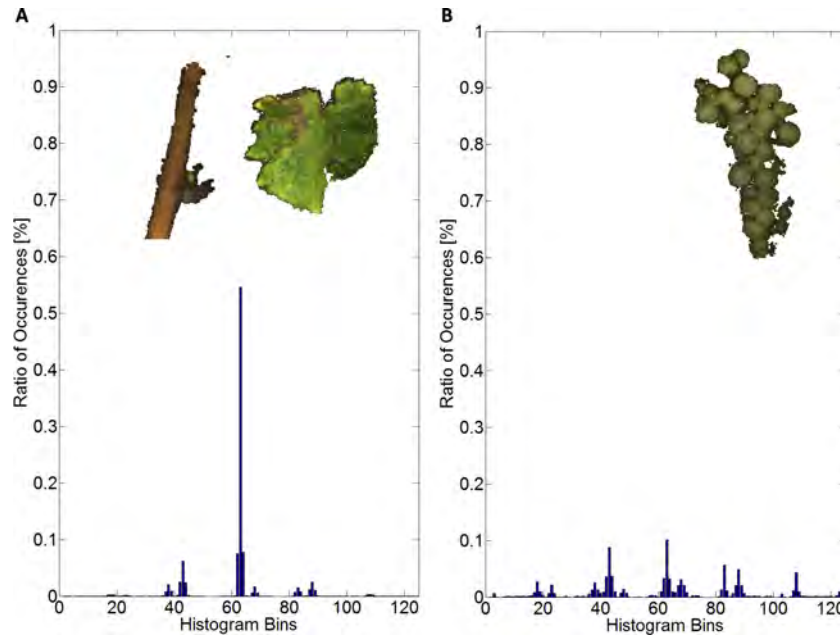


Abbildung 3.12: **A:** Typisches Surface-Feature-Histogramm der Klasse **Rest**. **B:** Typisches Surface Feature Histogramm der Klasse **Traube** für Weinreben im BBCH89-Reifestadium, aus Rose et al. [2016].

Ein Nachbarpunkt erhält damit maximal 50% Gewicht, wenn er genau dem Quellpunkt entspricht, wobei der Quellpunkt mind. 50% an Gewicht erhält. Der Quellpunkt erhält volle 100% Gewichtung, wenn die Distanz des Nachbarpunktes $d = r_h$ entspricht. Durch die Gewichtung und anschließende Mittelbildung der SPFHs

bleiben die SFHs im Wertebereich $[0..1]$ normiert, worin sie sich zusätzlich von den Arbeiten von Rusu [2009] unterscheiden. Die geschilderte Vorgehensweise wird für jeden Punkt $\mathbf{p}_i \in \mathcal{P}$ als Quellpunkt $\mathbf{p}$ wiederholt. Abbildung 3.12 zeigt typische SFH-Ausprägungen beider Klassen. Beide Histogramme unterscheiden sich in ihrer Ausprägung distinktiv voneinander.

3.2.2.7 Surface-Feature-Histogram-Color

Die SFH- und HSV-Deskriptoren werden zum SFHc-Deskriptor mit der Dimension $f = 125 + 3 = 128$ konkateniert:

$$\underset{(1 \times 128)}{\mathbf{sfhc}} = \begin{bmatrix} \underset{(1 \times 125)}{\mathbf{sfh}} & \underset{(1 \times 3)}{\mathbf{hsv}} \end{bmatrix}. \quad (3.20)$$

Der Wertebereich der SFHc-Deskriptoren ist wie der Wertebereich der SFc-Deskriptoren vollständig auf den Wertebereich $[0..1]$ normiert.

3.2.3 Überwachte vs unüberwachte Klassifikation

Die folgenden Ausführungen beziehen sich exemplarisch auf den SFHc-Deskriptor und die beiden in dieser Arbeit definierten Klassen **Rest** und **Traube**. Die Ausführungen gelten jedoch prinzipiell für jeden verwendeten Deskriptor und jede Art von Klassen.

Unüberwachte und überwachte Klassifikatoren sind die populärsten Methoden zur Klassifikation von Daten [Duda et al., 2012; Kotsiantis, 2007]. In unüberwachten Methoden, wie bspw. dem K-mean-clustering [Lloyd, 1982], werden die Deskriptorvektoren für die Daten zuerst berechnet und Klassen a-posteriori anhand eines Ähnlichkeitsmaßes ihrer Deskriptorvektoren definiert. Auf diese Weise können zwar unbekannte Klassen anhand ihrer Deskriptorvektoren entdeckt werden. Pflanzen variieren jedoch häufig in Geometrie und Farbe, bspw. nach Reifezustand oder Krankheitsbefall. So könnten Klassen generiert werden, die nicht mit tatsächlichen Pflanzenteilen korrespondieren oder die aus einer Komposition mehrerer, verschiedener Pflanzenteile bestehen, deren Deskriptorvektoren ähnlich sind.

Für diese Arbeit werden daher ausschließlich überwachte Klassifikatoren betrachtet, d.h. es werden anhand von Vorwissen 3D-Punkte $\mathbf{p}_i \in \mathcal{P}$ ausgewählt, die bzgl. Geometrie und Farbe repräsentative Instanzen der gewählten Klassen darstellen. Diese werden als sog. *Trainingsdaten* $\mathcal{T}$ mit M -Punkten $\mathbf{p}_i = [X, Y, Z]_m \in \mathbb{R}_3$, $m = 1, \dots, M$ bezeichnet. Die Gültigkeit der gewählten Trainingsdaten als Repräsentation aller Klasseninstanzen basiert auf der Annahme, dass Geometrie und Farbe der Instanzen einer Klasse relativ konstant über die betrachteten Daten verbleiben. Die manuelle Auswahl der Trainingsdaten ermöglicht die volle Kontrolle darüber, welche Geometrien und Farbwerte die einzelnen Klassen umfassen sollen. Dies ist insbesondere in der Phänotypisierung von Pflanzen von Bedeutung, da ein erfahrener Anwender um die Spanne möglicher Erscheinungsarten der Pflanze weiß. Er kann somit Klassen definieren, die mit den Pflanzenteilen von Interesse und ihren möglichen Erscheinungsformen korrespondieren.

3.2.4 Erstellung von Trainings- und Testdaten

Trainingsdaten beider Klassen werden manuell mittels der kommerziellen Software Geomagic Studio 12 (Raindrop Geomagic Inc., Morrisville, NC, USA) aus einer Punktwolke segmentiert. Jedem Punkt der Trainingsdaten wird anschließend sein *Klassenlabel* $c = \{1, 2\} \in \mathbb{N}$ zugewiesen, welches mit den Klassen **Rest** und **Traube** korrespondiert. In Abbildung 3.13 sind Beispiele ausgewählter Trainingsdaten für Riesling Pflanzen im BBCH89-Stadium abgebildet. Im Anschluss werden die SFHc-Deskriptorvektoren der Trainingsdaten berechnet. Folgende Notation wird festgelegt: Seien Trainingsdaten $\mathcal{T}$ aus M -Punkten und die mit den Trainingsdaten korrespondierenden SFHc-Deskriptorvektoren $(\mathbf{sfhc}_m, c_m)$, $m = 1, \dots, M$ gegeben, mit Klassenlabeln $c_m = \{1, 2\} \in \mathbb{N}$. Die Deskriptorvektoren werden in der $(M \times 128)$ -Matrix $\mathbf{D} = [\mathbf{sfhc}_1, \dots, \mathbf{sfhc}_M]$ zusammengefügt, die Klassenlabel im Vektor $\mathbf{c} = [c_1, \dots, c_M]$.

Überwachte Klassifikatoren nutzen die Deskriptorvektoren der Trainingsdaten $\mathcal{T}$ und das Vorwissen um die Klasse jedes Punktes der Trainingsdaten, um sog. *Klassifikationsmodelle*, d. h. Entscheidungsregeln, für die Klassifikation von Daten festzulegen. Die Berechnung eines Klassifikationsmodells bezeichnet man entsprechend als *Training*. Die Anzahl an Trainingsdaten für jede Klasse ist dabei ungefähr gleich groß, um eine Überanpassung des Modells an eine einzige Klasse zu verhindern. Die anhand des Klassifikationsmodells zu klassifizierenden Daten werden im Weiteren als *Testdaten* bezeichnet, welche in diesem Falle die rekonstruierte Punktwolke $\mathcal{P}$ darstellen. Zur Klassifikation der Testdaten $\mathcal{P}$ werden ebenfalls SFHc-Deskriptorvektoren für alle Punkte berechnet. Um Unabhängigkeit zwischen Trainings- und Testdaten zu gewährleisten, werden die Trainingsdaten aus einem separaten Datensatz ausgewählt.



Abbildung 3.13: Beispiele von Trainingsdaten. Die Farbe und Geometrie der Klassen **Rest** (Gerüstteile, Flatterbänder, Äste, Blätter) und **Traube** werden durch sie repräsentiert.

3.2.5 Überwachte Klassifikatoren

Ein Klassifikator weist anhand des Klassifikationsmodells und der Deskriptorvektoren jedes Punktes $\mathbf{p}_i \in \mathcal{P}$ jeden Punkt einer der c Klassen zu. In der Literatur sind eine Vielzahl überwachter Klassifikatoren zu finden [Jain et al., 2000; Kotsiantis, 2007]. Eine detaillierte Beschreibung aller existierenden Klassifikatoren liegt jedoch außerhalb der Reichweite dieser Arbeit. Für alle Klassifikationsaufgaben wird in dieser Arbeit die Offline-Implementation der Import-Vector-Machine (IVM) verwendet, wie sie in Roscher et al. [2012a] und Roscher [2012] beschrieben sind. Im Folgenden werden die überwachten Klassifikatoren *Random Forests* und *Support Vector Machine* als zwei der populärsten Klassifikatoren zusätzlich zu den Import Vector Machine kurz vorgestellt. Die Funktionsweise der IVM wird hingegen eingehend erläutert und die Wahl der Import Vector Machine als Klassifikator für diese Arbeit begründet.

3.2.5.1 Random Forests

Random Forests (RF) [Breiman, 2001] setzen sich aus einem Ensemble von *Decision Trees* (dt: Entscheidungsbäume) zusammen. Decision Trees (DT) sind hierarchisch konzipiert und bestehen aus Knoten, an denen logische Entscheidungsregeln anhand eines Deskriptoreintrages aus den Trainingsdaten festgelegt werden. Zum Training an einem Wurzelknoten mit einer ersten logischen Regeln startend, werden anhand einzelner Deskriptoreinträge an jedem Knoten neue logische Regeln hierarchisch und sequentiell hinzugefügt. Dies geschieht, bis kein neuer Knoten mehr entsteht und eine Entscheidung über die Klassenzugehörigkeit getroffen werden kann.

Das Klassifikationsmodell eines RF besteht aus einem Ensemble aus DTs, die jeweils einen zufällig gewählten

Satz an Deskriptoreinträgen heranziehen. Zur Klassifikation der Testdaten werden die einzelnen Klassifikationsergebnisse der DT zu einer endgültigen Klassenzuweisung nach dem Mehrheitsentscheid kombiniert. Wahrscheinlichkeiten der Klassenzugehörigkeit $W_p = [W_1, \dots, W_C]$ eines Punktes, wobei C die Anzahl der Klassen darstellt, können für jeden Punkt $\mathbf{p}_i \in \mathcal{P}$ bspw. aus der Häufigkeit der prädizierten Klasse jedes DT des RF in Relation zur Gesamtzahl der DT berechnet werden [Bostrom, 2007].

3.2.5.2 Support Vector Machine

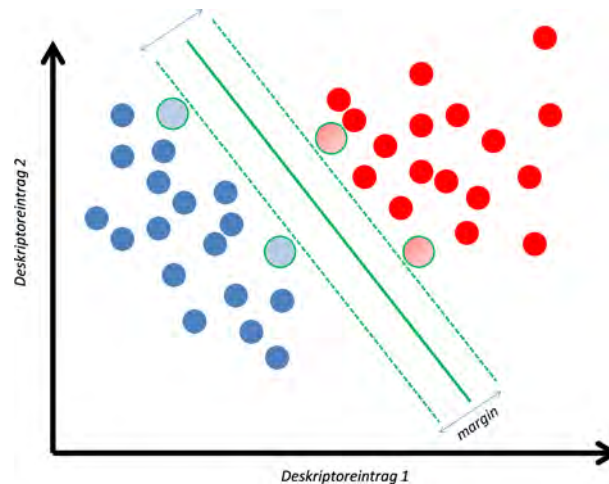


Abbildung 3.14: Ein exemplarischer 2D-Deskriptorraum mit einer linear trennenden Entscheidungsgrenze. Blau und rote Punkte symbolisieren die Deskriptorvektoren beider Klassen, Support Vectors sind grün umrandet, sowie farblich hervorgehoben. Die grüne geschlossene Linie stellt die diskriminierende Entscheidungsgrenze dar, die gestrichelten grünen Linien die Grenzen des *margin*.

Die *Support Vector Machine* (SVM) [Vapnik, 2013] trägt die einzelnen Deskriptorvektoren der Trainingsdaten als Punkte im sog. *Deskriptorraum* auf, welcher den f -dimensionalen Raum bezeichnet, den die Deskriptorvektoren aufspannen. Im Deskriptorraum wird mittels der Trainingsdaten eine lineare *Entscheidungsgrenze* als Klassifikationsmodell geschätzt, welche die Deskriptorvektoren der einzelnen Klassen so voneinander trennt, dass möglichst alle Instanzen der jeweiligen Klasse auf einer Seite der Entscheidungsgrenze liegen. Man spricht von einer *linearen Separation*.

Der Abstand der Entscheidungsgrenze zu den nächstgelegenen Deskriptorvektoren der einzelnen Klassen (engl. *margin*), den *Support Vectors*, soll maximiert werden, um eine möglichst große Generalisierungskraft der diskriminierenden Entscheidungsgrenze zu erreichen. Die Position der Support Vectors bestimmt die Position und Breite (*margin*) der Entscheidungsgrenze. Abbildung 3.14 illustriert einen exemplarischen 2D-Fall mit linear trennender Entscheidungsgrenze und den sie bildenden Support Vectors.

Oft sind die Daten zunächst nicht linear separierbar. An dieser Stelle wird der sog. *Kernel-Trick* [Aizerman, 1964] genutzt, um die Punkte der Trainingsdaten in den sog. *Kernelraum* zu transformieren und dort linear zu separieren. Eine Wahrscheinlichkeit der Klassenzugehörigkeit W_p jedes Punktes kann bspw. nach der Methode von [Platt, 1999] berechnet werden.

3.2.5.3 Import Vector Machine

Wie die SVM generiert die IVM ein diskriminatives Klassifikationsmodell basierend auf einer Entscheidungsgrenze zur Trennung der Klassen im Deskriptorraum. Sie nutzen dazu logistische Regression unter Anwendung des Kerneltricks für die Separation nicht-linear trennbarer Daten. Ergebnis der Klassifikation

sind daher A-posteriori-Wahrscheinlichkeiten für die Klassenzugehörigkeit jedes Punktes der Testdaten, wobei C die Anzahl der Klassen darstellt. IVMs sind hinsichtlich der Klassifikationsgenauigkeit gleichwertig wie oder höherwertiger als die SVM, klassifizieren jedoch schneller [Braun et al., 2011; Zhu and Hastie, 2005].

Die unbekannten Parameter des Klassifikationsmodells $\mathbf{a} = [\dots; \mathbf{a}_c; \dots]$ der C -Klassen werden durch Minimierung der Funktion

$$Q(\mathbf{a}) = Q_0(\mathbf{a}) + \frac{\lambda}{2} \mathbf{a}^T \mathbf{L} \mathbf{a}, \quad (3.21)$$

mit

$$Q_0(\mathbf{a}) = -\frac{1}{M} \sum_c \sum_m t_{mc} \log W_{mc}(\mathbf{a}). \quad (3.22)$$

Dabei ist $\mathbf{L}$ eine positiv definite symmetrische Matrix und λ ein positiver Regularisierungsparameter. Mit t_{mc} werden die Zielwahrscheinlichkeiten und mit W_{mc} die A-posteriori-Wahrscheinlichkeiten der Klassenzugehörigkeit zu den C -Klassen der zu klassifizierenden M -Punkte bezeichnet.

Für die Berechnung der Parameter des Klassifikationsmodells wird ein Subset $\mathcal{V} \in D$ mit V -Deskriptorvektoren $\mathbf{sfc}_v, v = 1, \dots, V$, den sog. *Import Vectors (IV)* verwendet. Mittels des Subsets wird das Klassifikationsmodell über alle Trainingsdaten approximiert, was die Komplexität der Berechnungen reduziert und Arbeitsspeicher einspart [Roscher et al., 2012b].

Die Deskriptorvektoren in $\mathcal{D}$ und $\mathcal{V}$ werden dazu zunächst mittels einer Kernelfunktion in den Kernelraum mit der Dimension V transformiert. Die Kernelfunktion lautet

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2} \frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\sigma^2}\right), \quad (3.23)$$

wobei $(\mathbf{x}, \mathbf{x}')$ jeweils Deskriptorvektoren derselben Dimension darstellen und σ den Kernelparameter bezeichnet. Es ergeben sich die $(M \times V)$ -Kernelmatrix $\mathbf{K}_V$, sowie die $(V \times V)$ -Regularisierungsmatrix $\mathbf{K}_R$ nach:

$$\mathbf{K}_V = k(\mathbf{sfc}_m, \mathbf{sfc}_v), \quad (3.24)$$

$$\mathbf{K}_R = k(\mathbf{sfc}_v, \mathbf{sfc}_{v'}), \quad (3.25)$$

mit $\mathbf{sfc}_m \in \mathcal{D}$ und $\mathbf{sfc}_v, \mathbf{sfc}_{v'} \in \mathcal{V}$. Die unbekannten Parameter des Klassifikationsmodells $\mathbf{a} = [\dots; \mathbf{a}_c; \dots]$ der C -Klassen werden für die Iteration t nach der *iterated reweighted least-squares (IRLS)*-Methode [Roscher, 2012] bestimmt. Die in $\mathcal{D}$ enthaltenen Deskriptorvektoren werden dabei in den Kernelraum transformiert:

$$\mathbf{a}_{c,(t)} = \left(\frac{1}{M} \mathbf{K}_V^T \mathbf{R}_c \mathbf{K}_V + \lambda \mathbf{K}_R\right)^{-1} \mathbf{K}_V^T \mathbf{R}_c \mathbf{z}_c \quad (3.26)$$

$$\mathbf{z}_c = \frac{1}{M} (\mathbf{K}_V \mathbf{a}_{c,(t-1)} + \mathbf{R}_c^{-1} (\mathbf{W}_c - \mathbf{t}_c)), \quad (3.27)$$

wobei $\mathbf{t}_c$ die Zugehörigkeit des Deskriptors $\mathbf{sfc}_m$ zur Klasse c indiziert. $\mathbf{R}_c$ ist eine Diagonalmatrix mit Einträgen $r_{mm} = W_m(1 - W_m)$ auf der Hauptdiagonalen, welche die Wahrscheinlichkeit der M -Punkte zur Klasse c enthält. Die A-posteriori-Wahrscheinlichkeiten der Klassenzugehörigkeit W_{mc} der zu klassifizierenden M -Punkte ergeben sich anschließend unter Nutzung der nun bekannten Parameter des Klassifikationsmodells über die Formel

$$W_{mc}(\mathbf{a}) = \frac{\exp(\mathbf{a}_c^T \mathbf{k}_m)}{\sum_{c'} \exp(\mathbf{a}_{c'}^T \mathbf{k}_m)}, \quad (3.28)$$

wobei $\mathbf{k}_m$ die Spalte mit Index m der Kernelmatrix $\mathbf{K}_V$ darstellt.

Das Subset $\mathcal{V}$ wird sukzessiv um zufällig ausgewählte IV erweitert oder verringert [Roscher et al., 2012b]. Im Gegensatz zu den Support Vectors der SVM, welche eine optimal diskriminierende Entscheidungsgrenze generieren, werden die IV dabei so ausgewählt, dass sie die A-posteriori-Wahrscheinlichkeit der Klassenzugehörigkeit W_{mc} für alle Trainingsdaten der Klasse c maximieren. Sie definieren daher nicht zwangsläufig die Position der Entscheidungsgrenze [Roscher et al., 2012a]. Abbildung 3.15 illustriert einen exemplarischen 2D-Fall mit nicht-linear trennender Entscheidungsgrenze und den sie bildenden Import Vectors.

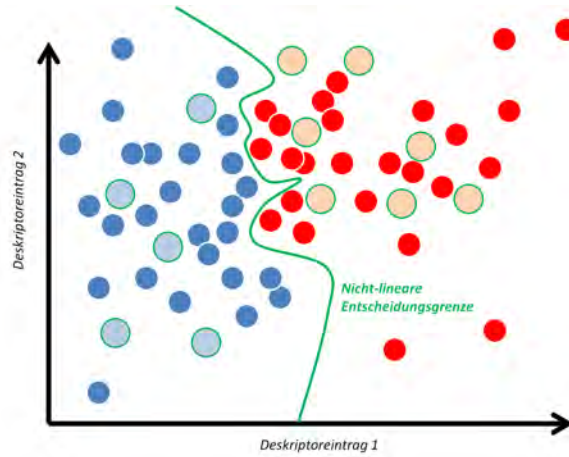


Abbildung 3.15: Ein exemplarischer 2D-Deskriptorraum mit einer nicht-linearen trennenden Entscheidungsgrenze. Blau und rote Punkte symbolisieren die Deskriptorvektoren beider Klassen, Import Vectors sind grün umrandet, sowie farblich hervorgehoben. Die grüne geschlossene Linie stellt die diskriminierende Entscheidungsgrenze dar.

Optimale Werte für den Kernparameter σ und den Regularisierungsparameters λ werden anhand der Trainingsdaten über Kreuzvalidierung vor der eigentlichen Klassifikation bestimmt. Für σ werden Werte im Bereich $0,25 - 11,3$, für λ Werte im Bereich $e^{-12} - e^{-3}$ untersucht. Die Klassifikation der Testdaten $\mathcal{P}$ erfolgt im Anschluss anhand des Klassifikationsmodells mit den beiden optimalen Parametern.

Die Klassifikation mit IVMs wird als *initiale Klassifikation* bezeichnet. Ergebnis der initialen Klassifikation sind die Wahrscheinlichkeiten der Klassenzugehörigkeit $W_p = [W_1, \dots, W_C]$ jedes Punktes $\mathbf{p}_i \in \mathcal{P}$. Ein Punkt wird derjenigen Klasse c zugewiesen, für welche die Wahrscheinlichkeit am höchsten ist:

$$\mathbf{p}_i \in \begin{cases} 1 & \max(W_p) = W_1 \\ 2 & \max(W_p) = W_2 \\ \dots & \\ c & \max(W_p) = W_C \end{cases} \quad (3.29)$$

3.2.5.4 Wahl der IVM als Klassifikator

Die Ergebnisse der initialen Klassifikation werden in einem weiteren Schritt mittels der Software *GCO* unter Nutzung der räumlichen Nachbarschaftsbeziehungen der 3D-Punkte und deren Wahrscheinlichkeiten der Klassenzugehörigkeit W_p verbessert (s. folgendes Kapitel). Eine gute Schätzung von W_p ist daher von elementarer Bedeutung für den Erfolg der Klassifikation.

Nach Chawla and Cieslak [2006]; Franc et al. [2011] oder Niculescu-Mizil and Caruana [2005] sind die Methoden zur Schätzung der Wahrscheinlichkeit der Klassenzugehörigkeit W_p im Falle der SVM und des RF jedoch nicht verlässlich. Die Import Vector Machine hingegen basiert auf einem probabilistischen Klassifikationsmodell durch logistische Regression. Das Ergebnis der Klassifikation eines Quellpunktes $\mathbf{p}$ ist daher eine *natürliche Schätzung* [Zhu and Hastie, 2005] der A-posteriori-Wahrscheinlichkeiten. Die IVM sind für diese Arbeit daher als Klassifikator vor den RF und SVM zu bevorzugen, da sie eine verlässlichere Methode zur Schätzung der Wahrscheinlichkeit der Klassenzugehörigkeit W_p darstellen.

3.2.6 Graph-Cut Software GCO

Im vorherigen Kapitel wurde beschrieben, wie jeder Punkt mittels der IVM klassifiziert wird. Ein Punkt wird dabei unabhängig von seinen Nachbarpunkten klassifiziert, d. h. ohne Informationen über die prädierte

Klassen seiner Nachbarpunkte einzubeziehen. Mit großer Wahrscheinlichkeit gehören benachbarte Punkte jedoch derselben Klasse an. In einer begrenzten Punktregion sei *Klassifikationsrauschen* definiert als Punkte einer divergierenden Klasse innerhalb der sonst weitestgehend homogen klassifizierten Punktregion einer anderen Klasse. In den folgenden Erläuterungen wird beispielhaft entsprechend der Klassen **Traube** und **Rest** von 2 Klassen ausgegangen.

In den initial klassifizierten Testdaten soll das Klassifikationsrauschen reduziert werden, indem alle benachbarten Punkte innerhalb einer begrenzten Punktregion der dort dominanten Klasse zugewiesen werden. Dazu wird die open-source software GCO (GCOptimization) [Boykov and Kolmogorov, 2004; Boykov et al., 2001; Delong, 2016; Delong et al., 2012; Kolmogorov and Zabini, 2004] verwendet, welche sowohl das Klassenlabel als auch die W_p der Punkte zur Homogenisierung nutzt.

Die Klassifikation wird dabei als Energiefunktion ausgedrückt. Die optimale Klassifikation ist diejenige, welche am wenigstens Energie verbraucht. Die aufzuwendende Energie $E(c)$ für die Klassifikation eines Punktes soll daher minimiert werden, um eine optimale Klassenzuweisung $\tilde{c} = \operatorname{argmin}_c E(c)$ zu erhalten. Die Energie $E(c)$ ergibt sich aus dem sog. *Datenterm* $D(c)$ und dem sog. *Glattheitsterm* $G(c)$. Es sei folgende Notation gegeben:

$$E(c) = D(c) + G(c) \quad (3.30)$$

mit

$$D(c) = \sum_{p \in \mathcal{P}} D_p(c_p) \quad (3.31)$$

$$G(c) = \sum_{pq \in \mathcal{P}^k} w_{pq} \cdot G_{pq}(c_p, c_q). \quad (3.32)$$

Mit $\mathbf{p}_i$ sei ein 3D-Punkt in $\mathcal{P}$ mit dem Klassenlabel c_p aus der initialen Klassifikation bezeichnet. $D(c)$ enthält den negativen Logarithmus der Wahrscheinlichkeiten der Klassenzugehörigkeit $W_p = [W_1 \ W_2]$ des Punktes $\mathbf{p}_i$, d. h. hohe Wahrscheinlichkeiten entsprechen einem geringen, niedrige Wahrscheinlichkeiten einem hohen Energieaufwand. Für $D(c)$ gilt:

$$D(c) = -\log(W_p), \quad (3.33)$$

wobei für alle Punkte in $\mathcal{P}$ berechnet wird.

Die Summenanteile im Glattheitsterm werden jeweils für eine lokale Nachbarschaftsregion $\mathcal{P}^k$ innerhalb eines gewählten Radius r_G für jeden Punkt $\mathbf{p}_i \in \mathcal{P}$ und seine Nachbarpunkte $\mathbf{q} \in \mathcal{P}^k$ mit den Klassenlabeln c_q formuliert. Der Gewichtungsfaktor w_{pq} ist hier ein frei wählbares Maß der Ähnlichkeit zwischen zwei benachbarten Punkten $(\mathbf{p}, \mathbf{q})$.

Zwischen den jeweiligen Punktepaaren $(\mathbf{p}, \mathbf{q})$ wird aufgrund ihrer räumlichen Nähe Glattheit, d. h. *Identität der Klassenlabel*, angenommen. Der lokale Energieaufwand G_{pq} , der für den Fall $c_p \neq c_q$ und einen Wechsel von einem Klassenlabel zum anderen erbracht werden muss, wird in einer symmetrischen Matrix formuliert. Es wird das *Potts Model* angenommen. Nach Boykov et al. [2001] gilt:

$$Potts(c_p, c_q) = \begin{cases} 0 & \text{für } c_p = c_q \\ 1 & \text{für } c_p \neq c_q \end{cases} \quad (3.34)$$

Da nur zwei Klassen prädiziert werden, besitzt G_{pq} die Dimension (2×2) . Der Energieaufwand für einen Wechsel der Klasse ist paarweise identisch, so dass gilt:

$$G_{pq} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (3.35)$$

Der Energieaufwand entspricht demnach $0 \cdot w_{pq}$, falls die Klassenlabel von p und q identisch sind, andernfalls $1 \cdot w_{pq}$. Grundsätzlich gilt, dass $D(c)$ die Energie global bestimmt, während $G(c)$ die lokale Energie enthält.

GCO interpretiert die 3D-Punkte in $\mathcal{P}$, sowie die Klassen $c = 1, 2$ als Knoten $\mathcal{V}$ mit Kanten $\mathcal{E}$, die jeweils zwei Knoten verbinden. Zusammen bilden sie den *Graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, dessen Kantenwerte den Werten aus dem Daten- bzw. Glattheitsterm entsprechen. Ein sog. *Schnitt* unterteilt die Datenpunkte des Graphen $\mathcal{G}$ durch Schneiden der Kanten so in die zwei Klassen **Rest** und **Traube**, dass gilt:

$$\mathcal{V} = \text{Rest} \cup \text{Traube} \quad (3.36)$$

$$1 \cap 2 = \emptyset \quad (3.37)$$

$$1 \in \text{Rest} \quad (3.38)$$

$$2 \in \text{Traube} \quad (3.39)$$

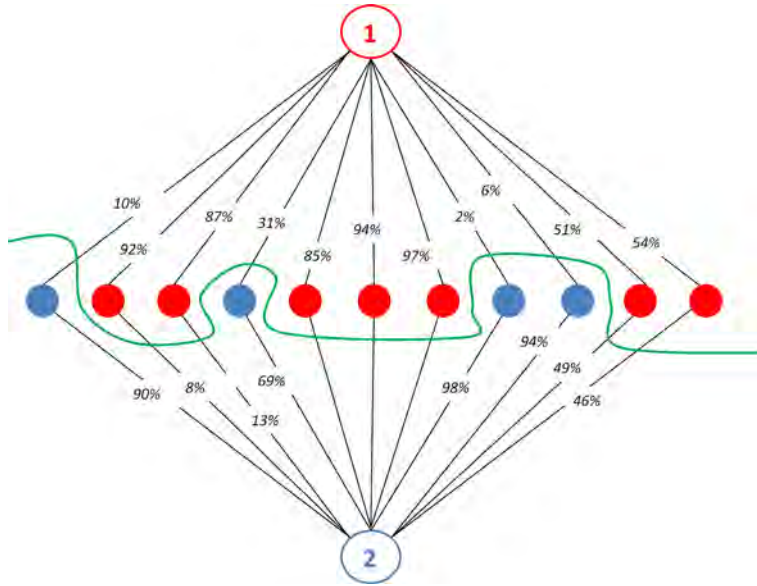


Abbildung 3.16: Ein exemplarischer 1D-Graph $\mathcal{G}$. Der Schnitt ist durch die grüne Linie dargestellt. Sein Verlauf entspricht der aufzuwendenden Energie für die Klassenzuweisung aus der initialen Klassifikation und trennt die Datenpunkte beider Klassen durch den Schnitt durch die verbindenden Kanten von Datenpunkt und Klassenlabel. Diese Energie für den Schnitt soll minimiert werden.

Abbildung 3.16 illustriert den Sachverhalt anhand eines 1D-Beispiels. Die zwei Klassenlabeln $c = 1, 2$ (roter bzw. blauer Kreis) und Datenpunkte (gefüllter rote bzw. blaue Kreise) bilden zusammen die Menge $\mathcal{V}$ der Knoten des Graphen. Die Datenpunkte wurden in der initialen Klassifikation jeweils einer Klasse zugewiesen, dargestellt durch ihre Farbe. Die schwarzen Linien stellen die Kanten zwischen den Klassenlabeln und den Datenpunkten dar. Die Werte der Kanten entsprechen den Einträgen in $D(c)$ (Es sind jedoch nur die Wahrscheinlichkeiten $W_p = [W_1 \ W_2]$ eingezeichnet).

Der dargestellte Schnitt (grüne Kurve) entspricht dem Schnitt nach der initialen Klassifikation. Die Summe der Werte der geschnittenen Kanten ist äquivalent zur aufzuwendenden Energie für die Unterteilung in die zwei Klassen. Die Energie für diesen Schnitt soll unter Bezug auf die Energie (Formel 3.30) minimiert werden. Die Minimierung der Energie entspricht dem Schnitt, welcher die Summe der Werte der verbindenden Kanten minimiert [Kolmogorov and Zabin, 2004].

GCO berücksichtigt bei der Berechnung des minimalen Schnitts zusätzlich die Nachbarschaftsrelationen der Datenpunkte. Abbildung 3.17 zeigt das Ergebnis nach der Neuuzuweisung der Klassenlabel für jeden Datenpunkt. Den Datenpunkten wurden neue Klassenlabel entsprechend des Schnitts durch die Kanten mit minimaler Energie (grüne Linie) für die Klassenzuweisung zugewiesen.

Im 3D-Fall geht GCO analog vor, die Nachbarn eines Punktes werden lediglich innerhalb einer Kugel mit Radius r_G um den Punkt bestimmt. Den 3D-Punkten der Punktwolke werden im Zuge der Energieminimierung

dabei neue Klassenlabel zugewiesen. Dadurch entstehen Punktregionen mit homogener Klassenzuweisung. Dieser zweite Schritt der Klassifikation mittels GCO wird als *Klassenglättung* bezeichnet.

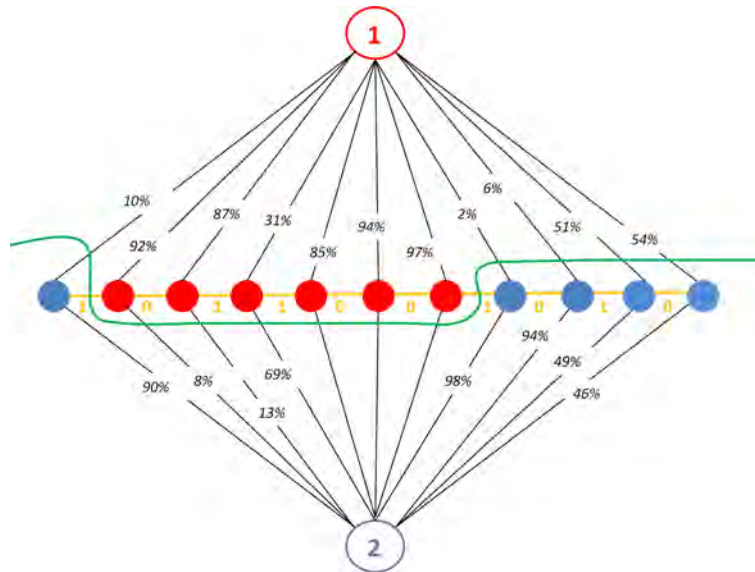


Abbildung 3.17: Der exemplarische 1D-Graph $\mathcal{G}$ nach der Energieminimierung mit GCO. Die orangefarbene Linie zwischen den Datenpunkten indiziert gegenseitige Nachbarschaften zwischen den Datenpunkten. Zusammen bilden schwarze und orangefarbene Kanten die Menge der Kanten $\mathcal{E}$ des Graphen. Die orangefarbenen Werte zwischen benachbarten Datenpunkten entsprechen dem Glattheitsterm nach dem Potts Model G_{pq} , welches abweichende Klassenlabel zwischen Nachbarn bestraft. Der Schnitt ist glatter, der Energieaufwand für die Klassenzuweisung wurde minimiert.

3.2.7 Zusammenfassung

Mehrere auf Geometrie und Farbe basierende Deskriptoren wurden vorgestellt. Um denjenigen Deskriptor zu identifizieren, mit dem das beste Klassifikationsergebnis erreicht werden kann, muss jeder Deskriptor untersucht werden. Dabei sind die jeweilig zu wählenden Parameter der Deskriptoren zu beachten, die großen Einfluss auf das Klassifikationsergebnis haben. Die geometrischen Deskriptoren hängen von der Wahl des Nachbarschaftsradius r_n und im Falle der SFH- und CSHOT-Deskriptoren zusätzlich von der Wahl des Histogrammradius r_H und der Anzahl b der Histogrammintervalle ab. Die Wahl der Radien ist nicht trivial und jede Kombination von Normalen- und Histogrammradius muss über Referenzdaten mit bekannter Klassenzuweisung hinsichtlich der Klassifikationsgenauigkeit evaluiert werden.

Der zu erwartende positive Effekt der Klassenglättung mittels GCO auf die Klassifikationsgenauigkeit muss ebenfalls anhand dieser Referenzdaten evaluiert werden. 3D-Punkte, die der Klasse **Traube** angehören, aber fälschlicherweise der Klasse **Rest** zugewiesen wurden, können nicht für die Schätzung der Ertragsparameter genutzt werden. Es ist daher zu prüfen, wie viele Trauben durch den Klassifikator nicht korrekt klassifiziert werden können.

Mehrere Klassifikatoren wurden vorgestellt. Die IVM wurde aufgrund ihrer probabilistischen Eigenschaften als Klassifikator für diese Arbeit anderen Klassifikatoren vorgezogen.

4. Praxisbeispiel 1: Voruntersuchungen mit dem PHENObot

Im Folgenden werden Voruntersuchungen beschrieben, die auf Daten der Saison 2015 basieren und mit dem PHENObot am Julius-Kühn-Institut (JKI) in Siebeldingen, Deutschland, erhoben wurden. Im Zuge der Voruntersuchungen wird die Auswertemethode zur Lösung aller drei Teilaufgaben der Datenerhebung, Datenklassifikation und Quantifikation eingehend behandelt. Bei der Lösung der drei Teilaufgaben kommen sowohl kommerzielle, frei verfügbare sowie eigens für NoViSys entwickelte Hard- und Software-Komponenten zum Tragen. Die einzelnen Komponenten werden detailliert erläutert und Voruntersuchungen zu den einzelnen Komponenten beschrieben. Die Voruntersuchungen erfolgen dabei in der Reihenfolge der drei Teilaufgaben. Um ihre gegenseitige Verknüpfung auch im Lesefluss darzustellen, erfolgen die Erläuterungen der Komponenten innerhalb der Teilschritte der Auswertemethode ebenfalls in Reihenfolge ihrer Anwendung. Einige vorgestellte Ergebnisse wurden bereits in Rose et al. [2016] veröffentlicht und sind entsprechend als solche gekennzeichnet.

Zunächst wird der PHENObot und die grundlegende Vorgehensweise bei der Bildaufnahme für die Rekonstruktion der Anbaureihen mittels SfM/MVS beschrieben. Im Anschluss werden die Ergebnisse der Rekonstruktion der Anbaureihen in beiden Reifestadien beider Erziehungssysteme vorgestellt. Darauf erfolgt eine Evaluierung der im vorherigen Kapitel vorgestellten Deskriptoren für die Klassifikation von Punktwolken der Anbaureihen, bei der der beste Deskriptor für die Klassifikation der Anbaureihen identifiziert wird. Die Funktionsweise des Algorithmus *Berryfinder* zur Detektion von Beeren wird darauffolgend erläutert und die Genauigkeit der Beerendetektion sowohl anhand von Referenzdaten aus dem Labor, als auch mittels Daten aus dem Feld untersucht. Im letzten Teil werden für 5m lange Ausschnitte dieser vier Anbaureihen die Anzahl der für die Kameras sichtbaren Trauben und Beeren, sowie die Beerengröße anhand des *Berryfinders* in Kombination mit einem 3D-Connected Components Verfahren bestimmt und die Ergebnisse mit manuell erzeugten Referenzwerten verglichen.

4.1 Datenerhebung mit dem PHENObot

4.1.1 PHENObot

Der PHENObot wurde ursprünglich im Zuge der Projekte PHENOVines und CROPSENSE [PHENOVines, 2017; Project CROP.SENSE, 2017] entwickelt und besitzt bereits die meisten wesentlichen Komponenten, die für die Datenerhebung relevant sind. Er stand bereits unmittelbar zu Beginn des Projektes zur Verfügung und wurde daher bis zur Realisierung des Phenoliners für erste Felduntersuchungen mit dem Ziel eines *proof-of-concepts* und zur Erhebung erster Datensätze genutzt, auf denen die Algorithmenentwicklung stattfand. Da seine Entwicklung jedoch nicht im Kontext von NoViSys geschah, sei für eine detailliertere Beschreibung der Funktionsweise des PHENObots und weitere mit ihm durchgeführte Untersuchungen an dieser Stelle auf die Ausführungen in Kicherer [2015] und Kicherer et al. [2015a] verwiesen.

Der PHENObot ist ein Raupenfahrzeug von 1,2m Länge, 1,10m Breit und 1,50m Höhe. Die für NoViSys relevanten Komponenten des PHENObots sind:

- Raupenfahrzeug mit geringer Bewegungsvibration zur Aufnahme scharfe Bilder als Sensorträger.
- RTK-GPS für die Georeferenzierung der Bilder (Trimble® SPS852, Geo Systems GmbH, Jena, Deutschland).
- Getriggert LED-Beleuchtungsrahmen (ODLW300, Smart vision lights, Muskegon, MI, USA) mit einer Vielzahl individueller Lichtquellen.

- $1 \times 5\text{MP}$ RGB Industriekamera (AVT GT-2450C, $u = 2448 \times v = 2050\text{pixel}$; horizontale aktive Sensorlänge 8mm , vertikale Sensorlänge 7mm ; Objektiv: Schneider KMP-IR CINEGON 8mm Brennweite).
- Schnelle Erhebung und Speicherung von Bilddaten mit eigener Software (IggGeotagger).

Abbildung 4.1 zeigt den PHENObot im Feldeinsatz. Die Bewegung des PHENObots wird manuell über eine Fernsteuerung kontrolliert. Die maximale Fahrgeschwindigkeit beträgt 6km/h . Das RTK-GPS befindet sich auf dem Rahmen des PHENObots und übertagt die Weinreben um wenige Dezimeter.

Fünf Kameras (drei Monochromkameras, eine NIR-Kamera, eine RGB-Kamera) sind an einem beweglichen Rahmen montiert, mit dem sich der vertikale und senkrechte Abstand der Kameras zur Weinreihe adaptieren lässt. Für die Untersuchungen werden lediglich die Bilder der in der Mitte montierten RGB-Kamera verwendet. Die Beleuchtungseinheit umrahmt in rechtwinkliger Anordnung die Kameras. Die flächenhaft angebrachten LEDs sowie die Diffuserplatten vor den LEDs erzeugen eine gleichmäßige Ausleuchtung der Weinreben.

Kameras, Beleuchtungseinheit, RTK-GPS und Recheneinheit sind über ein Ethernet Netzwerk miteinander verbunden. Die Aufnahme und Speicherung von Bildern geschieht über die Software IggGeotagger. Das manuelle Betätigen eines Knopfes löst die Kamera aus, die wiederum die Beleuchtungseinheit triggert. Die Bilder werden auf der Recheneinheit abgelegt und die GPS-Position zum Zeitpunkt der Bildaufnahme im Exif-Header der Bilder gespeichert.



Abbildung 4.1: Der PHENObot, aus Rose et al. [2016]

4.1.2 Bildaufnahme mit dem PHENObot

4.1.2.1 Betrachtete Weinreben

Insgesamt werden jeweils eine 25m lange TS- und eine MSS-Anbaureihe im Reifestadium BBCH75 und BBCH89 erfasst. Im Reifestadium BBCH75 handelt es sich um eine Riesling Anbaureihe nach dem MSS-Erziehungssystem und eine Chardonnay Reihe nach dem TS-Erziehungssystem. In diesem Reifestadium haben beide Sorten die für dieses Reifestadium dieselbe typische, hellgrüne Beerenfarbe, wobei die typische Charakteristik beider Erziehungssysteme hinsichtlich Blattwerk und Trauben bereits ausgebildet ist. Im Reifestadium BBCH89 wird je eine Riesling-Anbaureihe nach dem TS- bzw. MSS-Erziehungssystem betrachtet. Die Beeren besitzen eine identische, grün-bräunliche Farbe.

4.1.2.2 Vorgehensweise

Die Bildaufnahme mit dem PHENObot fand nach Einbruch der Dunkelheit statt, um homogene Lichtverhältnisse zu gewährleisten. Die Bildaufnahme wird durch manuelles Drücken eines Knopfes aus der Fahrt heraus ausgelöst. Der Beleuchtungsrahmen erhellt die Weinreben gleichmäßig und durchgehend, ohne Schlagschatten zu erzeugen. Für die Bildaufnahme bewegte sich der PHENObot im Korridor zwischen zwei Anbaureihen, bei einer Geschwindigkeit von etwa $0,3\text{km/h}$. Die Kamera ist entlang der Fahrtrichtung nach rechts auf die Anbaureihe und parallel zum Boden ausgerichtet, der Abstand H zwischen Kamera und Pflanzen beträgt ca. 750mm . In den MSS-Anbaureihen ragten stellenweise lange Äste in den Fahrweg des PHENObots und hätten dadurch die Kamera verdecken können. Diese Äste wurden möglichst umfahren oder der Ast durch einen Nutzer während des Vorbeifahrens ferngehalten. Dieses Phänomen trat bei den Anbaureihen im TS-Erziehungssystem nur selten auf.

4.1.2.3 Einnahme multipler Perspektiven, Gewährleistung der Bildüberlappung und Zeitaufwand

Für die Rekonstruktion der Anbaureihen mittels Multi-View-Stereo werden Bilder aus mehreren Perspektiven mit einer ausreichend hohen Bildüberlappung entlang der Horizontalen und der Vertikalen benötigt (s. Kapitel 3.1.2). Der horizontale Abstand zweier aufeinanderfolgender Bilder sei mit $base_h[\text{mm}]$ bezeichnet, der vertikale Abstand entsprechend mit $base_v[\text{mm}]$. Die Berechnung der Bildüberlappung entlang der Horizontalen und der Vertikalen bei den Abständen $base_h$ bzw. $base_v$ wird im Weiteren geschildert.

Zunächst muss die Strecke S_O berechnet werden, welche der Bildsensor entlang seiner Bildachsen im Objektraum erfasst. Diese wird entsprechend der Geometrie der Zentralprojektion einer Kamera nach dem Strahlensatz berechnet. Bestimmende Parameter sind die physische Länge des Kamerasensors S entlang einer der Bildachsen, die Kamerakonstante $F[\text{mm}]$ ($\approx$ Brennweite der Kameralinse) und der Abstand zum Objekt H . Abbildung 4.2 stellt die Zentralprojektion und das Verhältnis der Parameter dar.

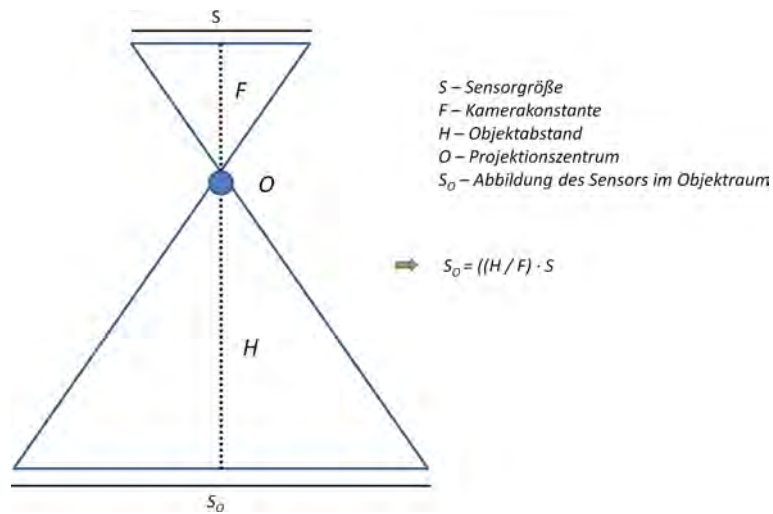


Abbildung 4.2: Die Berechnung der projizierten Strecke des Bildsensors im Objektraum erfolgt über den Strahlensatz. Der blaue Punkt stellt das Projektionszentrum der Kameralinse dar, durch das alle Lichtstrahlen gemäß des Prinzips der Zentralprojektion fallen.

S_O berechnet sich folgendermaßen:

$$S_O = \frac{H}{F} \cdot S. \quad (4.1)$$

Die horizontale aktive Sensorlänge S des PHENObots beträgt $8mm$, die vertikale Sensorlänge $7mm$, die Kamerakonstante F beträgt $8mm$. Nach Formel 4.1 beträgt die horizontal projizierte Strecke damit $S_O = 750mm$, die vertikal projizierte Strecke S_O beträgt $656mm$. Die horizontale Bildüberlappung o_h beträgt bei einer horizontalen S_O dann

$$o_h = 1 - \frac{base_h}{S_O} [\%]. \quad (4.2)$$

Die horizontale Verschiebung der Kamera in eine neue Perspektive wird durch die Bewegung der Sensorplattform durch die Anbaureihen erreicht. Die Bildaufnahme wird etwa alle $10 - 15cm$ ausgelöst. Der Abstand $base_h$ von rund $10 - 15cm$ sichert nach Formel 4.2 eine horizontale Bildüberlappung o_h von $80 - 90\%$. Auf diese Weise wird die gesamte Anbaureihe einmal bei einer initialen Kamerahöhe von $110cm$ über dem Boden vollständig durchfahren und bildlich erfasst.

Da der PHENObot lediglich eine einzige RGB-Kamera besitzt, muss die Anbaureihe zwei weitere Male mit jeweils höheren Kamerapositionen über dem Boden durchfahren werden, um multiple Perspektiven auch in der Höhe einzunehmen und möglichst die gesamte Ausbreitung des Blattwerks entlang der Vertikalen ($110 - 130cm$ vertikale Ausdehnung beim MSS-Erziehungssystem) bildlich zu erfassen (Abbildung 4.3). Der dazu notwendige vertikale Abstand $base_v$ zwischen den drei Kamerahöhen wird im Folgenden berechnet.

Damit die zwei weiteren Kamerahöhen für die Rekonstruktion verwendet werden können, muss auch vertikal eine gewisse Bildüberlappung o_v gewährleistet sein. Die Bildüberlappung sollte 60% nicht unterschreiten [Tsingas, 1992]. Um eine ausreichend hohe Bildüberlappung bei der komplexen Morphologie der Weinreben zu gewährleisten, wird eine vertikale Bildüberlappung von $o_v = 70\%$ angestrebt. Bei einer angestrebten Überlappung o_v berechnet sich der notwendige vertikale Abstand $base_v$ nach

$$base_v = (1 - o_v) \cdot S_O, \quad (4.3)$$

wobei S_O die vertikal projizierte Strecke des Sensors im Objektraum ist. Bei drei Kameras im Abstand von $base_v$ ergibt sich insgesamt eine vertikale Bildabdeckung S_v über

$$S_v = S_O + 2 \cdot base_v. \quad (4.4)$$

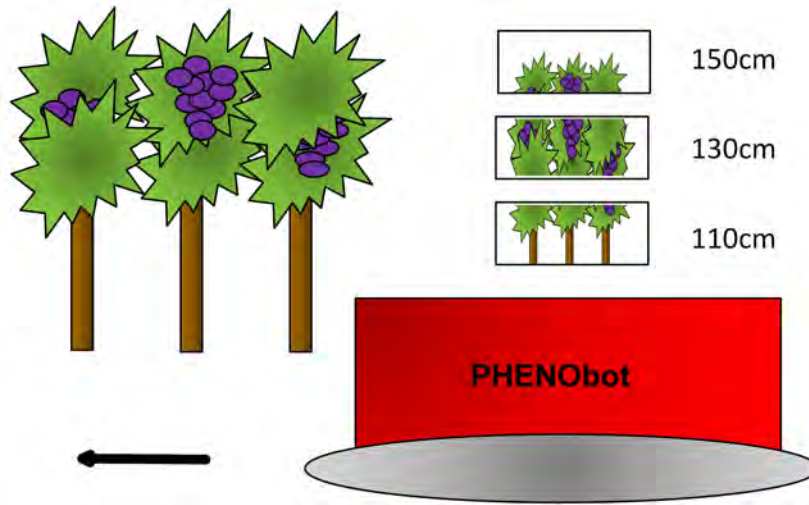


Abbildung 4.3: Datenerhebung mit dem PHENObot. Die Anbaureihe wird insgesamt dreimal durchfahren, um Bilder aus drei verschiedenen Höhen aufzunehmen, aus Rose et al. [2016].

Bei $o_v = 70\%$ und der vertikal projizierten Strecke $S_O = 656mm$ ergibt sich nach Formel 4.3 $base_v \approx 20cm$. Die zwei weiteren Kamerahöhen betragen damit $130cm$ und $150cm$. Insgesamt wird mit allen drei

Kamerahöhen vertikal nach Formel 4.4 eine Strecke $S_v = 1056\text{mm}$ abgedeckt. Das etwa $110 - 130\text{cm}$ hohe Blattwerk im MSS-Erziehungssystem wurde aufgrund der vertikalen Bildabdeckung von $S_v \approx 105\text{cm}$ vertikal zum Großteil, aber nicht vollständig erfasst. Trauben der MSS-Anbaureihen können sich daher evtl. stets außerhalb der Bildabdeckung befinden. Im TS-Erziehungssystem hingegen wachsen die Trauben dicht über dem Boden und in primär einem Höhenniveau. Es kann daher davon ausgegangen werden, dass die vertikale Bildabdeckung S_v für TS-Anbaureihen ausreichend ist, um vertikal alle sichtbaren Trauben zu erfassen.

Die Bildaufnahme mit dem PHENObot benötigt pro Anbaureihe aufgrund der drei notwendigen Durchfahrten ca. 15min . Dabei ergibt sich die Dauer nicht bloß aus den reinen Durchfahrten, sondern auch aus Wendemanövern am Ende der Anbaureihe, die Rückfahrt und erneute Wendung, um die Anbaureihe erneut durchfahren zu können. Pro Anbaureihe wurden etwa 700 Bilder (ca. 10GB) aufgenommen.

4.1.2.4 Ground Sampling Distance

Ein weiterer wichtiger Faktor bei der Rekonstruktion der Anbaureihen ist die *Ground Sampling Distance* (*GSD*). Sie ist ein Maß für das räumliche Auflösungsvermögen von Kameras, d. h., wie groß die tatsächliche Strecke im Objektraum ist, die ein einziges Pixel abbildet. Da die Punktwolken aus Bildern rekonstruiert werden, ist die GSD auch ein Maß für das geometrische Auflösungsvermögen der SfM/MVS Methode. Im Folgenden sei die *GSD* exemplarisch anhand der horizontal ausgerichteten Bildachse des Kamerasensors berechnet. Sie ist für beide Bildachsen identisch, da ein Pixel quadratisch ist. Die *GSD* ergibt sich, indem S_O durch die Pixelauflösung des Sensors u dividiert wird:

$$GSD = \frac{S_O}{u}. \quad (4.5)$$

Der PHENObot besitzt eine horizontale Pixelauflösung u von 2448pel . Es ergibt sich nach Formel 4.5 für die GSD eines einzigen Pixels

$$GSD = \frac{750\text{mm}}{2448} \approx 0,3\text{mm}. \quad (4.6)$$

Die GSD der Kamera des PHENObots ist mit $0,3\text{mm}$ somit hoch genug, um die Krümmung der nur wenige Millimeter großen Beeren mit ausreichend hoher Auflösung zu erfassen. Der Abstand H jedes Objektpunktes zur Kamera wird in Weinreihen jedoch variieren. Die berechnete GSD ist daher nur als Durchschnittswert zu betrachten.

4.1.3 3D-Rekonstruktion der Anbaureihen

4.1.3.1 Rekonstruktionsdauer und Größe der Punktwolke



Abbildung 4.4: Die Punktwolke einer 25m langen TS-Anbaureihe, aus Rose et al. [2016].

Mit Pix4DMapper wurden anhand der aufgenommenen Bilder vollständige, farbige 3D-Punktwolken der Anbaureihen in etwa 8-9 Stunden auf einem PC mit 8GB DDR-RAM Arbeitsspeicher rekonstruiert. Die Punktwolke einer vollständigen Anbaureihe besteht durchschnittlich aus etwa $70 - 100$ Millionen 3D-Punkten, je nach Erziehungssystem und Reifestadium. Abbildung 4.4 zeigt eine vollständige rekonstruierte, 25m lange Punktwolke einer TS-Anbaureihe.

4.1.3.2 Genauigkeit der Rekonstruktion

Der durchschnittliche Rückprojektionsfehler über alle Punkte der jeweiligen Punktwolken in beiden Erziehungssystemen und Reifestadien (s. Kapitel 3.1) betrug zwischen $0,1\text{pel}$ und $0,2\text{pel}$. Die Bedingung $d \ll 1\text{pel}$ ist damit für alle Punktwolken erfüllt. Die Punkt-zu-Punkt-Distanz der Punktwolken lag bei $0,3 - 0,5\text{mm}$. Die durchschnittliche räumliche Punktdichte der Punktwolke entspricht damit ungefähr der GSD von $0,3\text{mm}$.

Der Editor von Pix4DMapper gibt Auskunft über die Anzahl der Bilder, in denen ein rekonstruierter Objektpunkt beobachtet wurde, sowie über die Varianz aus allen individuellen Rückprojektionsfehlern des Objektpunktes in diese Bilder.

Die Häufigkeit, mit der ein Objektpunkt einer Traube durchschnittlich in den Bildern beobachtet wird, kann als ein Indikator dafür gesehen werden, wie wahrscheinlich es für häufig verdeckte Trauben und andere Objekte ist, die Mindestanzahl von drei homologen Bildpunkten für die Rekonstruktion zu erreichen. Je häufiger die Weinreben beobachtet werden, desto wahrscheinlicher wird diese Mindestanzahl erreicht.

Der durchschnittliche Rückprojektionsfehler d eines einzelnen Objektpunktes gibt die Unsicherheit seiner Rekonstruktion an und hängt sowohl von der Häufigkeit der Beobachtungen als auch den Schnittwinkeln zwischen den Projektionsstrahlen ab (s. Kapitel 3, Abb. 3.4).

Die Häufigkeit der Beobachtungen und der Rückprojektionsfehler eignen sich daher zum Vergleich der Rekonstruktion beider Erziehungssysteme. Es ist zu erwarten, dass die Unsicherheit der Rekonstruktion, sowie die Häufigkeit der Beobachtung eines Objektpunktes für TS- und MSS-Erziehungssysteme aufgrund ihrer unterschiedlichen Erscheinungsform und dem Unterschied in der Dichte des Blattwerks unterschiedlich sein werden.

Um die durchschnittliche Häufigkeit der Beobachtungen und den durchschnittlichen Rückprojektionsfehler von Trauben in beiden Erziehungssystemen festzustellen wurden im Pix4DMapper Editor manuell 11 3D-Punkte auf insgesamt vier Trauben ausgesucht und pro 3D-Punkt 1) die Varianz des Rückprojektionsfehlers aus allen den Punkt beobachtenden Bildern und 2) die Anzahl der Bilder erfasst, in denen er beobachtet wurde. Um die unterschiedlichen Wachstumspositionen der Trauben zu berücksichtigen wurden Trauben ausgesucht, die unterschiedlich tief im Blattwerk lagen. Die 11 Punkte wurden gleichmäßig auf der gesamten Traube ausgewählt. Insgesamt werden so jeweils 44 Werte für die Varianz sowie die Anzahl der Bilder erhalten. Diese werden anschließend gemittelt. Tabelle 4.1 listet die mittlere Anzahl der Beobachtungen eines Punktes, sowie die mittlere Varianz des Rückprojektionsfehlers der 44 einzelnen 3D-Punkte für beide Erziehungssysteme auf.

	$\overline{\text{Bilder}}$	$\overline{\sigma_d^2} [\text{pel}]$
TS-Anbaureihe	21	0,2
MSS-Anbaureihe	18	0,7

Tabelle 4.1: Mittlere Anzahl der Beobachtungen und mittlere Varianz aus 44 3D-Punkten beim PHENObot.

In TS-Anbaureihen wird ein Objektpunkt auf einer Traube durchschnittlich in 21 Bildern beobachtet, in MSS-Anbaureihen in 18. Die mittlere Varianz des Rückprojektionsfehlers ist für die TS-Anbaureihe niedriger als für die MSS-Anbaureihe. Beide Umstände resultieren vermutlich aus den unterschiedlichen Dichten des Blattwerks in beiden Erziehungssystemen. Trauben können aufgrund des dichteren Blattwerks im MSS-Erziehungssystem seltener und nur von annähernd gleich bleibenden Perspektiven aus beobachtet werden. Die Schnittwinkel α der Projektionsstrahlen verbleiben daher durchschnittlich kleiner als im Falle des TS-Erziehungssystems, wo aufgrund des geringeren Verdeckungsgrades größere Schnittwinkel der Projektionsstrahlen aus einer Vielzahl unterschiedlicher Perspektiven möglich sind. Es lässt sich konstatieren, dass Trauben im TS-Erziehungssystem eher und genauer rekonstruiert werden, als im MSS-Erziehungssystem. Die Rekonstruktion von Trauben im MSS-Erziehungssystem ist aufgrund des dichten Blattwerks hingegen schwieriger.

Die innere Genauigkeit und die geometrische Auflösung der Rekonstruktion beider Erziehungssysteme kann anhand der Ergebnisse als hoch genug aufgefasst werden, die Trauben und die runde Oberfläche der kleinmaßstäbigen Beeren von wenigen Millimetern Durchmesser korrekt darzustellen.

4.1.3.3 Georeferenzierung der Bilder

Die Bilder des PHENObots werden über die aktuelle GPS-Position zum Aufnahmezeitpunkt georeferenziert. Pix4DMapper nutzt die GPS-Positionen der Bilder als Näherungswerte für die Kameraposition in der Bündelausgleichung. Jedoch entspricht die physische Position der Kamera nicht der physischen Position der GPS-Antenne auf dem PHENObot. Eine zentimetergenaue Transformation der GPS-Position der GPS-Antenne auf die tatsächliche Position der vertikal und in der Lage beweglichen Kamera findet im Igg-Geotagger nicht statt. In Pix4Dmapper wird daher eine Standardabweichung der Kameraposition von 30cm angegeben.

In der Bündelausgleichung führt dies dazu, dass die Position der Kameras primär auf den Beobachtungen der homologen Punkte und SfM beruht und die absolute Lage der Punktwolke um mehrere Dezimeter verschoben sein kann. Die durchschnittliche Residue zwischen der in der Bündelausgleichung geschätzten Kameraposition und gemessener GPS-Position der Kamera betrug nach der von Pix4DMapper ausgegebenen Statistik im Mittel über alle vier Anbaureihen in der Lage $1 - 3\text{cm}$ für den Rechtswert, $2 - 7\text{cm}$ für den Hochwert und $10 - 30\text{cm}$ für die Höhe. Hohe Residuen in der Positionsbestimmung der Kameras außerhalb der spezifizierten Unsicherheit der GPS-Positionsbestimmung von 2cm sind darauf zurückzuführen, dass insgesamt drei Kamerahöhen von $110 - 150\text{cm}$ verwendet wurden, woraus die im Vergleich zu der Lageposition deutlich höheren Residuen in der Höhe resultieren. Abbildung 4.5 zeigt eine rekonstruierte Anbaureihe sowie die initialen und ausgeglichenen GPS-Positionen der Kameras.

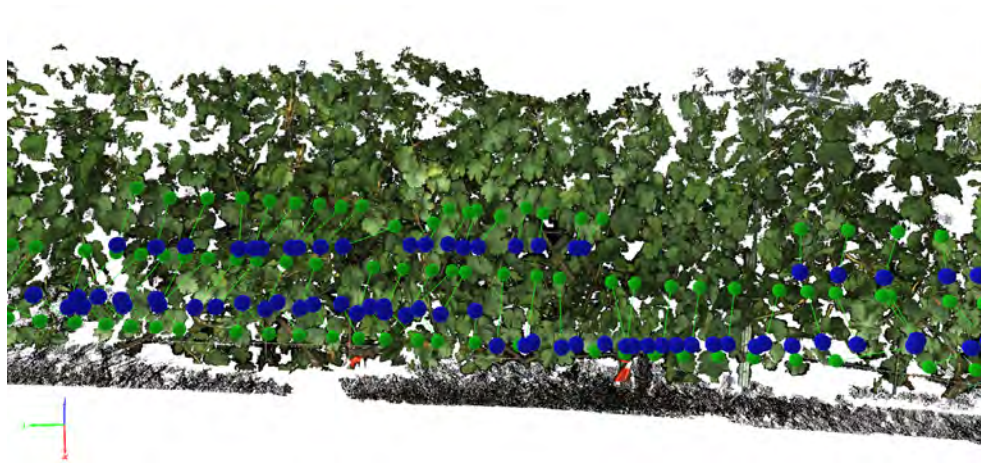


Abbildung 4.5: Ausschnitt aus einer rekonstruierten MSS-Anbaureihe. Blaue Kugeln indizieren initiale GPS-Positionen der Kameras, grüne Kugeln geschätzte Kamerapositionen $\mathbf{p}_{cam}$ nach der Bündelblockausgleichung. Grüne Verbindungsgeraden zwischen den Kugeln indizieren das Ausmaß der Positionsresiduen, aus Rose et al. [2016].

4.1.4 Vorverarbeitung der Punktwolken

Durch Abweichungen in den Kameraparametern und Kamerapositionen oder homogene Bildinhalte kann es zur Rekonstruktion von 3D-Punkten an falschen Positionen oder zur Generierung von 3D-Punkten kommen, die keinem tatsächlichen Objektpunkt entsprechen (s. Kapitel 3.1). Die Punktwolke ist zudem mit zufälligen Abweichungen behaftet und besitzt mit 3D-Punkten im Abstand von $0,3 - 0,5\text{mm}$ eine hohe Punktdichte, die auf ein für den Arbeitsspeicher geeignetes Maß reduziert werden muss. Die Punktwolke wird daher für

die spätere Klassifikation und Quantifikation vorverarbeitet.

Das folgende Kapitel beschreibt die Vorgehensweise bei der Vorverarbeitung der Punktwolken für die Datenklassifikation. Die Vorverarbeitung umfasst vier Schritte: 1) Entfernung des Hintergrundes, 2) Entfernung von Ausreißern, 3) Reduktion zufälliger Abweichungen, 4) Reduktion der Punktdichte.

4.1.4.1 Entfernung des Hintergrundes

Bei der Hintergrundentfernung werden zwei Phänomene berücksichtigt. Das erste Phänomen sind Punktregionen, die nicht mit tatsächlich existenten Objekten korrespondieren. Diese entstehen durch die Triangulation homogen gefärbter Bildpunkte des nachtdunklen Hintergrunds. Die triangulierten 3D-Punkte befinden sich zumeist in der Nähe von Objekten mit heterogener Färbung. Durch die Nähe zu den Objekten mit heterogener Färbung können auf der Epipolarlinie homogen gefärbte Punkte des Hintergrunds als homolog identifiziert und trianguliert werden, die nicht tatsächlich miteinander korrespondieren. Diese Punkte äußern sich in dünnen, unstrukturierten Punktregionen ohne erkennbare Geometrie und entsprechen keinem tatsächlichen Objekt an dieser Position.

Ein zweites Phänomen sind 3D-Punkte des Hintergrunds, die in den Bildern an Objekträndern liegen und dort auch trianguliert worden sind. Sie verdicken die tatsächlichen Objektränder und verfälschen damit die Objektgeometrie. Zur Entfernung des Hintergrundes werden die RGB-Werte aller Punkte in den *HSV-Raum* (*Hue-Saturation-Value*) transformiert (s. Kapitel 3.2.2.1). Das RGB-Modell stellt Farbe durch die Addition aller RGB-Kanalwerte dar. Die Identifikation bestimmter Farben oder Helligkeitswerte anhand von Wertebereichen für die RGB-Werte ist daher kompliziert. Die Transformation in den HSV-Raum ermöglicht dagegen die lineare Unterscheidung von Farbe, Sättigung und Helligkeit der Punkte anhand einzelner Schwellwerte für die HSV-Werte.



Abbildung 4.6: **A:** Ausschnitt aus einer Punktwolke. An den Rändern der Objekte sind Punkte in der schwarzen Farbe des Hintergrunds rekonstruiert worden, welche die Objektgeometrie verfälschen und keinen tatsächlichen Objektpunkt entsprechen (Beispiele rot markiert). In der Nähe heterogen gefärbter Objekte entstehen des Weiteren ungeordnete, dünne Punktregionen des Hintergrunds (Beispiel blau markiert). **B:** Der Ausschnitt nach der Entfernung des Hintergrundes. Die falsch triangulierten Punkte wurden größtenteils entfernt. Kleinere Regionen dieser Punkte verbleiben jedoch weiterhin in der Punktwolke. Objekte tiefer im Blattwerk der Weinrebe, die nicht ausreichend ausgeleuchtet wurden, können unter Umständen ebenfalls entfernt werden.

Als Punkte des Hintergrunds gelten Punkte, die einen Value-Wert (Helligkeit) von $V = 0,12$ nicht erreichen. Dieser Wert wurde empirisch anhand von Punkten des Hintergrunds durch Mittelbildung ihrer Value-Werte bestimmt. Punkte unterhalb dieses Wertes werden unabhängig von ihren Hue- und Saturation-Werten gelöscht (Abbildung 4.6). Einige wenige Objekte, die tiefer im Blattwerk liegen und nicht genügend ausgeleuchtet werden konnten, werden unter Umständen ebenfalls gelöscht. Diese schwach ausgeleuchteten Objekte wurden jedoch aufgrund ihrer geringen Ausleuchtung ohnehin häufig nicht exakt rekonstruiert.

Alle folgenden Vorverarbeitungsschritte werden mittels der frei verfügbaren Software CloudCompare (CloudC) durchgeführt [Danielgm, 2017].

4.1.4.2 Entfernung von Ausreißern

Ausreißer werden mittels des *Statistical Outlier Removal (SOR)-Filters* entfernt. Zunächst werden der Abstand d eines Punktes zu einer nutzerdefinierten Anzahl an nächstliegenden Nachbarn, sowie die Standardabweichung σ_d der Abstände berechnet. Der Abstand des Punktes wird anschließend mit dem durchschnittlichen Punktabstand d_{mean} unter allen Nachbarn verglichen. Es wird ein Schwellwert d_{max} für die zulässige Distanz über

$$d_{max} = d_{mean} + f \cdot \sigma_d \quad (4.7)$$

festgelegt. Dabei ist f ein nutzerdefinierter Multiplikator für die Standardabweichung. Eine Nachbarschaftszahl von 100 Punkten und ein Faktor $f = 2$ hat sich dabei durch empirisches Ausprobieren und visuellen Vergleich als optimal erwiesen. Einzelne Ausreißer von Punkten, die an falscher Position rekonstruiert wurden, werden auf diese Weise gelöscht.

4.1.4.3 Reduktion zufälliger Abweichungen

Zur Klassifikation mit den geometrischen Deskriptoren müssen Normalenvektoren zu den durch die 3D-Punkte abgebildeten Geometrien (s. Kapitel 3.2.2.6) berechnet werden. Zufällige Abweichungen der 3D-Punkte von der Soll-Geometrie der Oberflächen führen jedoch zur Berechnung von Normalenvektoren, die nicht dem Krümmungsverlauf der tatsächlichen Oberflächengeometrie entsprechen (Abbildung 4.7). Zur Unterscheidung von Trauben und anderen Bestandteilen ist es jedoch wichtig, die Geometrie der Oberfläche möglichst exakt und damit unterscheidbar zu rekonstruieren. Zufällige Abweichungen in der Punktwolke werden mit dem *Point Cloud Library (PCL)* Plugin [Penasa, 2017] *Moving-Least-Squares (MLS)* [Alexa et al., 2003] reduziert. Dieser Vorgang wird folgend als *Glättung* bezeichnet.



Abbildung 4.7: Grüne Quadrate: 3D-Punkte einer beispielhaften Ebene. Blaue Kreise mit Pfeilen: Normalenvektoren zu beispielhaften 3D-Punkten der Ebene. **A:** Zufällige Abweichungen der 3D-Punkte von der Soll-Geometrie der Ebene resultieren in Normalenvektoren, deren Ausrichtung nicht dem tatsächlichen Krümmungsverlauf der Ebene entspricht. **B:** Die Glättung der Oberfläche der Ebene resultiert in parallel zueinander ausgerichteten Normalenvektoren, die dem Krümmungsverlauf der Ebene folgen.

Beim MLS-Verfahren werden 3D-Punkte einer Oberfläche lokal durch Schätzung einer Ebene geglättet. Dazu werden zunächst die Nachbarpunkte in einem nutzerdefinierten Glättungsradius r_s um einen Quellpunkt herum bestimmt. Mittels der Nachbarpunkte wird eine Ebene über den Bereich der Nachbarpunkte approximiert. Die Distanz jedes Nachbarpunktes zur lokalen Ebene wird über eine Polynomfunktion mit einem nutzerdefinierten Grad n approximiert. Dies wird mit jedem Punkt der Punktwolke als Quellpunkt wiederholt und resultiert in glatten Oberflächen.

Die Punktwolken werden mit einem Glättungsradius $r_s = 5mm$ für die BBCH89 und $r_s = 3mm$ für die BBCH75 Reifestadien und einem Polynomgrad von zwei nach dem MLS-Verfahren geglättet. Diese Werte haben sich durch visuellen Vergleich als optimal erwiesen. Zufällige Abweichungen werden ausreichend

reduziert und die typische Kugelgeometrie der Beeren beibehalten, welche die Trauben geometrisch unterscheidbar von lokal ebenen Blättern und zylinderförmigen Ästen macht. Für detaillierte Ausführungen zur MLS-Methode sei auf [Alexa et al., 2003] verwiesen.

4.1.4.4 Reduktion der Punktdichte

Die geglättete Punktwolke wird zuletzt auf einen minimalen Punktabstand von 1mm ausgedünnt. Diese Punktdichte hat sich bereits bei Paulus et al. [2014a] als geeignet für die Klassifikation erwiesen. Sie erhält geometrische Unterschiede zwischen den Trauben und anderen Objekten der Weinreihe und entlastet den Arbeitsspeicher bei den weiteren Verarbeitungsschritten.

4.1.5 Detailgrad der Punktwolken

Abbildung 4.8 zeigt exemplarisch ein RGB-Bild aus der TS-Anbaureihe im BBCH89 Reifestadium (A), die rekonstruierte und verarbeitete Punktwolke der Szene (B), sowie die Glattheit und den erreichten Detailgrad anhand einer Traube der TS-Erziehung (C und D). Alle Objekte sind durch die Entfernung des Hintergrund scharf konturiert. Von Blättern verdeckte Trauben konnten durch die multiplen Perspektiven rekonstruiert werden. Die Traubenseiten wurden ebenfalls bis zu einem gewissen Grad rekonstruiert. Erhebungen individueller Beeren sind deutlich erkennbar und weisen eine glatte, runde Oberflächengeometrie auf.

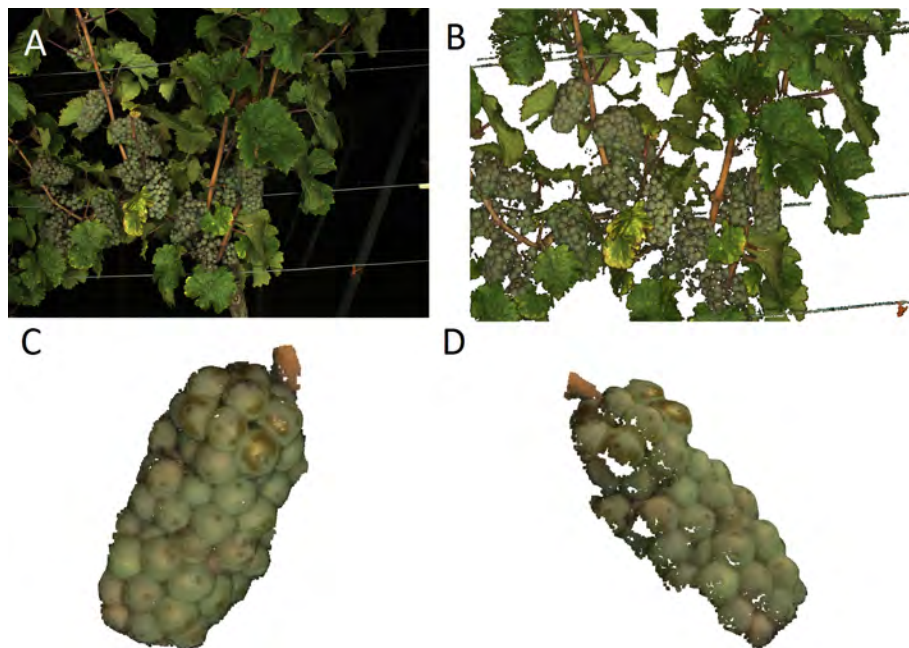


Abbildung 4.8: **A:** RGB-Bild einer TS-Anbaureihe im BBCH89 Reifestadium. **B:** Bereinigte und geglättete Punktwolke der selben Szene. **C:** Frontale Betrachtung einer Traube der Szene. **D:** Seitliche Betrachtung derselben Traube. (**B-D** wurden zur Visualisierung noch nicht ausgedünnt).

Abbildung 4.9 illustriert die selben Ansichten für einen Ausschnitt der MSS-Anbaureihe im BBCH75 Reifestadium. Hier wird die größere Dichte des Blattwerks deutlich. Die Rekonstruktion der im Vergleich zum BBCH89 Reifestadium deutlich kleineren Beeren erweist sich als schwieriger. Erhebungen individueller Beeren sind zwar auch hier erkennbar. Die eigentlich runde Oberflächengeometrie der Beeren kann jedoch nicht immer korrekt rekonstruiert werden. Beeren weisen teilweise eine abgeflachte Geometrie auf, die nicht dem tatsächlichen Radius der Beeren entspricht. Neben ihres kleineren Durchmessers ist dies auf Lichtreflexionen

größerer Intensität auf der Beerenmitte zurückzuführen. Diese könnten aufgrund der in diesem Reifestadium noch nicht stark ausgeprägten Wachsschicht der Beeren entstehen, die in späteren Reifestadien starke Reflexionen verhindert (s. Abbildung 4.8). Pixel auf der Beerenmitte besitzen durch die Reflexion sehr ähnliche Intensitätswerte und können daher bei der Verdichtung der Punktwolke während der Rekonstruktion nicht voneinander unterschieden werden (s. Kapitel 3.1.2 zu SfM/MVS). Diese Punkte werden an falschen Positionen rekonstruiert, wodurch die Kugelform der Beere flacher wird.

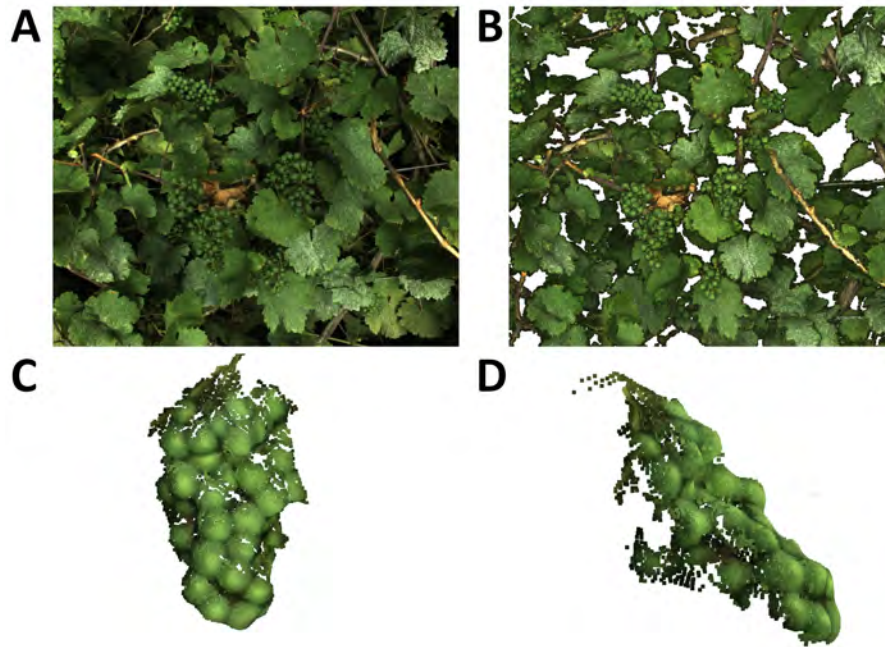


Abbildung 4.9: **A:** RGB-Bild einer MSS-Anbaureihe im BBCH75 Reifestadium. **B:** Bereinigte und geglättete Punktwolke der selben Szene. **C:** Frontale Betrachtung einer Traube im BBCH75 Reifestadium. **D:** Seitliche Betrachtung derselben Traube. (C-D wurden zur Visualisierung noch nicht ausgedünnt).

4.1.6 Zusammenfassung

Der PHENObot ist bedingt für die Ertragsschätzung unter Feldbedingungen geeignet. Er besitzt eine hochwertige Kamera und ein Beleuchtungssystem, welches für die Ausleuchtung von Weinreben optimal ist. Der mit dem Bildauslöser getriggerte Beleuchtungsrahmen ermöglicht eine vollständige, gleichmäßige und diffuse Ausleuchtung des abgebildeten Bereichs der Anbaureihe ohne direkte Reflexionen oder Schlagschatten. Die gleichmäßige Ausleuchtung wird u.a. dadurch erreicht, dass der Beleuchtungsrahmen fix um die Kamera montiert ist und sich bei einer Veränderung der Kameraposition mit bewegt. Unabhängig von der Kameraposition ist somit eine optimale Ausleuchtung gegeben. Dadurch werden farbechte und kontraststarke Bilder mit hoher Bildschärfe erzeugt.

Ein weiterer Vorteil des PHENObots ist seine Flexibilität und geringe Größe, die Anpassungen der Kameraausrichtung auf die Situation in der Anbaureihe wie bspw. herausragende Äste ermöglichen. Weinreben und Sensorik sind jedoch nicht abgeschirmt, wodurch homogene Lichtverhältnisse unabhängig von der Tageszeit nicht zu erreichen sind. Die Einsatzzeit des PHENObots nach Anbruch der Dunkelheit erzeugt zwar homogene Lichtverhältnisse, schränkt jedoch die Nutzungszeit ein.

Ein Nachteil ist, dass der PHENObot lediglich eine RGB-Kamera besitzt, weshalb zur Simulation eines vertikal angeordneten MKS insgesamt drei Durchfahrten durch die Reihe nötig sind. Die Erfassung einer Anbaureihe von 25m in drei Kamerahöhen benötigt daher etwa 15 Minuten. Der PHENObot ist somit *nicht* für die Erfassung von Anbaureihen im Hochdurchsatz geeignet.

Es konnte demonstriert werden, dass Multi-View-Stereo, gekoppelt mit georeferenzierten Bildern, eine geeignete Methode zur Rekonstruktion des Traubenbestandes der Anbaureihen ist. Die Position der GPS-Antennen wird jedoch nicht zentimetergenau auf die Position der Kamera transformiert. Bei entsprechend hoch angegebener Standardabweichung der Georeferenzierung beruht die Positionsbestimmung der Kameras daher primär auf SfM und der Bündelblockausgleichung. Durch die Georeferenzierung der Bilder wird direkt eine metrisch skalierte Punktwolke erhalten. Die Geometrie der kleinmaßstäbigen Beeren sowohl im BBCH89, wie im BBCH75 Reifestadium kann ausreichend genau rekonstruiert werden. Individuelle Beeren heben sich distinktiv von anderen Beeren ab und können für Kugelschätzungen genutzt werden.

4.2 Datenklassifikation

In diesem Kapitel werden zunächst die in Kapitel 3 vorgestellten geometrischen, farblichen und geometrisch-farblichen Deskriptoren hinsichtlich ihrer Eignung zur Klassifikation der Anbaureihen in beiden Reifestadien BBCH75 und BBCH89 untersucht. Ziel ist die Identifikation desjenigen Deskriptors, mit dem das beste Klassifikationsergebnis erzeugt werden kann und der für die weiteren Untersuchungen für die Klassifikation verwendet wird. Bei den Untersuchungen wird außerdem die im vorherigen Kapitel formulierte Idee untersucht, ob die Kombination geometrischer und farblicher Deskriptoren einen robusteren Deskriptor für die Klassifikation von Weinreben darstellt. Des Weiteren wird der Einfluss der Klassenglättung mittels GCO untersucht und eine Strategie zur Bestimmung optimaler Parameter für den verwendeten Deskriptor festgelegt.

Nach diesen vorausgehenden Untersuchungen werden pro Erziehungssystem und Reifestadium jeweils 5m lange Ausschnitte der Anbaureihen klassifiziert. Die als der Klasse **Traube** zugehörig identifizierten 3D-Punkte dienen im Kapitel 4.3 *Quantifikation* als Basis für die Detektion der sichtbaren Beeren und Trauben.

4.2.1 Evaluierung der Deskriptoren

Zur Identifikation des für die Klassifikation besten Deskriptors werden Testdaten $\mathcal{P}^t$ aus einer Anbaureihe entnommen und jeder Deskriptor einzeln zur Klassifikation der Testdaten herangezogen. Die Maße des Ausschnitts mit den Testdaten betragen rund $40 \times 40 \times 50 \text{ cm}$.



Abbildung 4.10: Die Untersuchung der Deskriptoren basiert auf den hier abgebildeten Testdaten $\mathcal{P}^t$. Der Ausschnitt besitzt die Ausmaße von rund $40 \times 40 \times 50 \text{ cm}$ und entstammt der Riesling Anbaureihe nach der MSS-Erziehung im Reifestadium BBCH89.

Zur Bewertung des Klassifikationserfolges der Deskriptoren werden die Metriken *Accuracy*, *Recall* und *Precision* verwendet. Recall (Vollständigkeit) und Precision (Präzision) werden über die Formeln 4.8 und 4.9

berechnet. Sie werden für jede Klasse separat berechnet und benötigen Referenzdaten der tatsächlichen Klassen der klassifizierten 3D-Punkte. Die tatsächlichen Klassen werden vorab von einem Anwender visuell bestimmt und jedem Punkt manuell zugewiesen. Die Metriken sind ein Maß für die Identität von tatsächlicher Klasse und durch Klassifikation zugewiesener Klasse eines 3D-Punktes.

Der Recall beantwortet folgende Frage: Wie viele Punkte einer Klasse, beispielsweise 1, wurden durch Klassifikation auch wirklich der Klasse 1 zugewiesen? Er gibt somit an, wie vollständig die Klassifikation Punkte einer Klasse bestimmt hat.

$$Recall = \frac{\#TruePositives}{\#TruePositives + \#FalseNegative} [\%], \quad (4.8)$$

wobei *True Positives* die Anzahl der Übereinstimmungen zwischen zugewiesener und tatsächlicher Klasse eines Datenpunktes bedeuten. Die Anzahl der *False Negatives* gibt an, wie viele Datenpunkte der Klasse 1 fälschlicherweise einer anderen Klasse, bspw. der Klasse 2, zugewiesen wurden. Die Summe der True Positives und False Negatives ist folglich die Gesamtanzahl der Klasse 1.

Die Precision ist die komplementäre Metrik zum Recall. Sie gibt an, wie viele der Datenpunkte, die durch automatische Klassifikation der Klasse 1 zugewiesen wurden, auch tatsächlich der Klasse 1 angehören. Sie wird nach folgender Formel berechnet:

$$Precision = \frac{\#TruePositives}{\#TruePositives + \#FalsePositives} [\%], \quad (4.9)$$

wobei ein *False Positive* einen Datenpunkt bezeichnet, welcher der Klasse 1 zugewiesen wurde, aber tatsächlich einer anderen Klasse angehört. Die *Accuracy* ist die Multiplikation von Recall und Precision:

$$Accuracy = Recall \times Precision [\%]. \quad (4.10)$$

Die Accuracy ist ein Maß für die Gesamtgenauigkeit der Klassifikation. Eine ideale Klassifikation hätte demnach einen Recall von 100% und eine Precision von 100%, mit einer Accuracy von 100%. Dann würden alle Datenpunkte der Klasse 1 detektiert und alle Punkte, die der Klasse 1 zugewiesen wurden, entsprächen auch tatsächlich der Klasse 1. Ein Recall von 80% bedeutete, dass 80% der Datenpunkte, die der Klasse 1 tatsächlich angehören, in den Punkten, die der Klasse 1 zugewiesen wurden, enthalten sind. Eine Precision von 60% bedeutete folglich, dass von allen der Klasse 1 zugewiesenen Punkten, 60% auch tatsächlich der Klasse 1 angehören. Im Umkehrschluss sind 40% aller der Klasse 1 zugewiesenen Datenpunkte in Wahrheit einer Klasse $\neq 1$ angehörig.

Alle künftigen Angaben dieser Metriken beziehen sich *ausschließlich* auf die Ergebnisse der Klasse **Traube**.

4.2.1.1 Vorgehensweise zur Identifikation des besten Deskriptors

Weinreben bestehen primär aus lokal ebenen Blättern, zylindrischen Ästen unterschiedlicher Dicke und dicht zusammen liegenden Beeren mit runder Oberflächengeometrie. Weiter in der Punktwolke vorhandene Objekte sind Drähte, Stützpfeiler und Markierungsmaterial. Deskriptoren müssen in der Lage sein, diese unterschiedlichen Objekte derart zu quantifizieren, dass distinktive Ausprägungen der Deskriptorvektoren beider Klassen entstehen, anhand derer beide Klasse unterschieden werden können. Jeder der vorgestellten Deskriptoren wird einzeln untersucht.

Die geometrischen Deskriptoren erzeugen unterschiedliche Deskriptorvektoren, je nachdem, wie die Berechnungsparameter r_n , r_H und b gewählt werden. Für jeden der geometrischen Deskriptoren müssen daher Berechnungsparameter bestimmt werden, die Accuracy, Recall und Precision maximieren. Die SF-Deskriptorvektoren hängen von der Wahl des Nachbarschaftsradius r_n für die Normalenberechnung ab, die SFH- und CSHOT-Deskriptorvektoren neben r_n zusätzlich von der Wahl des Histogrammradius r_H und der Anzahl b der Histogrammintervalle. Der HSV-Deskriptor hingegen wird direkt aus den RGB-Werten nach der in Kapitel 3.2.2.1 beschriebenen Vorgehensweise bestimmt und ist damit für gegebene RGB-Werte eines

Punktes stets *eindeutig* berechenbar. Optimale Berechnungsparameter müssen für ihn vorab nicht bestimmt werden.

Der Nachbarschaftsradius r_n bestimmt die Größe der zur Normalenberechnung betrachteten lokalen Geometrie um den Punkt $\mathbf{p}$. Da die Ausrichtung des Normalenvektors $\mathbf{n}_p$ von der Form der von r_n umschlossenen Geometrie abhängt, bestimmt r_n die Sensibilität der Deskriptoren hinsichtlich Krümmungsänderungen in der lokal betrachteten Geometrie.

Der Histogrammradius r_H bestimmt die Anzahl betrachteter Nachbarpunkte, die zur Berechnung der jeweiligen Einträge der Deskriptorvektoren anhand der vorab berechneten Normalenvektoren herangezogen werden. SFH und CSHOT hängen damit zusätzlich von der von r_H umschlossenen Geometrie ab. Je vielfältiger diese von r_H umschlossenen Geometrien werden, desto vielfältiger werden die berechneten Histogramme, wobei die Häufigkeit der Belegung jedes Histogrammintervalls von der relativen Häufigkeit der Geometrien in der betrachten Punktregion und der Anzahl der Histogrammintervalle b bestimmt wird.

Die Anzahl b der Histogrammintervalle haben nach Behley et al. [2012] keinen signifikanten Einfluss auf die Klassifikation, falls sie groß genug gewählt werden. Dieser Parameter wird daher in dieser Arbeit fix gehalten und orientiert sich an bisherigen Veröffentlichungen. Im Falle der SFH wird $b = 5$ [Paulus et al., 2013] und $b = 11$ bzw. $b = 31$ für den SHOT- bzw. CIELab-Anteil im CSHOT-Deskriptor [Tombari et al., 2011] gewählt.

Zur Bestimmung der optimalen Werte für r_n und r_H müssen eine Vielzahl von Werten für beide Radien zur Berechnung der geometrischen Deskriptorvektoren getestet werden. Dazu werden feste Werte für r_n und r_H innerhalb eines vorgegebenen Wertebereichs definiert. Folgendes wird festgelegt:

- Im Falle der SF-Deskriptoren werden für r_n Werte im Bereich $1 - 14mm$ mit $1mm$ Schrittweiten getestet:
 $r_n = [1, 2, 3, \dots, 13, 14]mm$.
- Im Falle der SFH- und CSHOT-Deskriptoren werden für r_n Werte im Bereich $1 - 5mm$ mit $1mm$ Schrittweiten und für r_H Werte im Bereich $7 - 42mm$ mit $5mm$ Schrittweiten getestet:
 $r_n = [1, 2, \dots, 5]mm$ und $r_H = [7, 12, \dots, 37, 42]mm$.
 Es werden *alle möglichen Kombinationen* der festgelegten Werte für r_n und r_H untersucht. Dabei gilt stets $r_H \geq r_n$.

Durch die Wahl der Wertebereiche werden bewusst Radien abgedeckt, die von wenigen Millimetern bis mehrere Zentimeter reichen können. Dadurch wird eine Vielzahl möglicher Geometrien abgedeckt. Auf diese Weise soll herausgefunden werden, in welcher geometrischer Größenordnung die optimalen Radien liegen.

Zur Identifikation des besten Deskriptors muss die Klassifikation mit allen vorgestellten Deskriptoren und pro geometrischen Deskriptor mit allen oben definierten Parameterwerten einzeln durchgeführt und das Klassifikationsergebnis anhand von Accuracy, Recall und Precision bewertet werden. Abbildung 4.11 stellt die Vorgehensweise in einem Flussdiagramm dar. Als bester Deskriptor gilt derjenige, mit dem unter allen Deskriptoren die höchsten Werte für Recall und Precision erreicht werden können. Um letztere zu berechnen, wird jedem Punkt der Testdaten $\mathcal{P}^t$ (Abbildung 4.10) manuell das Klassenlabel $c = 1, 2$ zugewiesen, welches entsprechend mit den Klassen **Rest** und **Traube** korrespondiert.

Des Weiteren werden wie in Kapitel 3.2 beschrieben Trainingsdaten T für beide Klassen **Rest** und **Traube** aus 3D-Punkten $\notin \mathcal{P}^t$ erstellt. Für jeden 3D-Punkt in $\mathcal{P}^t$ und T wird der jeweilig untersuchte Deskriptor herangezogen und die Deskriptorvektoren für alle Punkte berechnet.

Anschließend wird ein Klassifikationsmodell mittels der Deskriptorvektoren von T und den IVM generiert und mittels des Klassifikationsmodells die Testdaten $\mathcal{P}^t$ anhand ihrer Deskriptorvektoren des jeweilig untersuchten Deskriptors klassifiziert. Dies geschieht für den HSV-Deskriptor und alle geometrischen Deskriptoren mit allen oben definierten Werten und Kombinationen von r_n und r_H . Auf eine Klassenglättung mit GCO wird an dieser Stelle noch verzichtet, um den Deskriptor mit der größten eigenen Unterscheidungskraft identifizieren zu können.

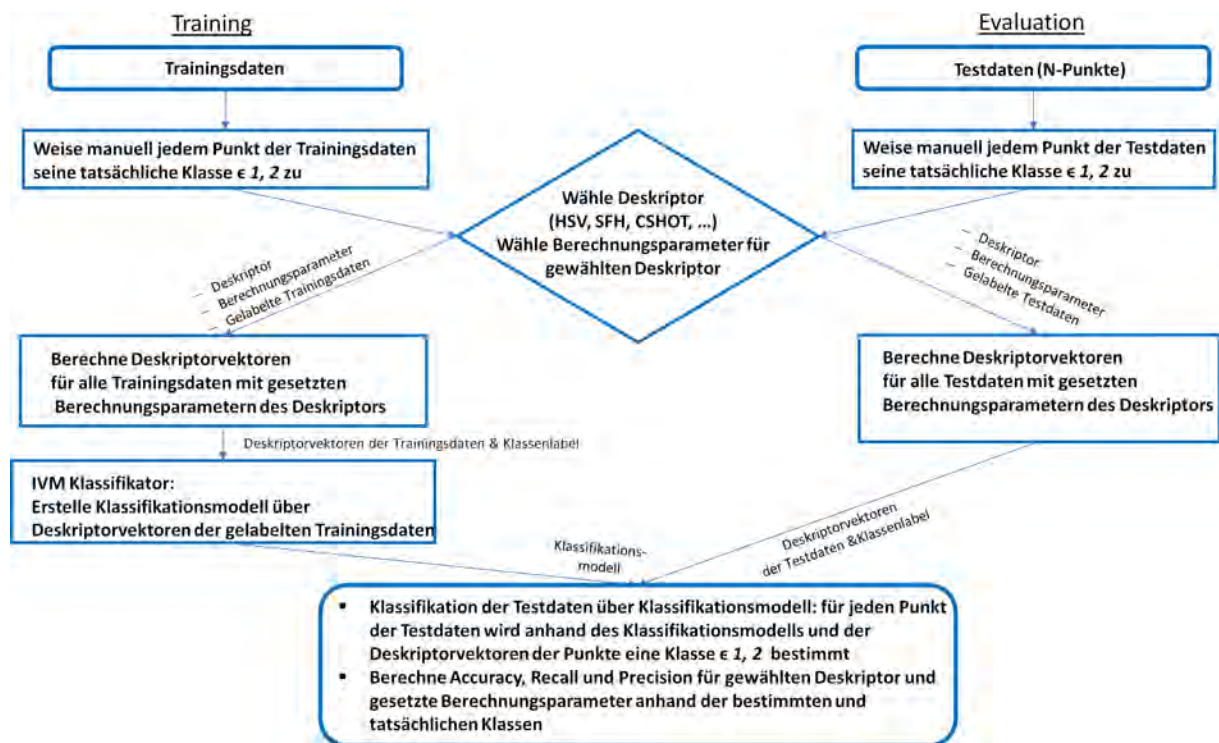


Abbildung 4.11: Flussdiagramm der Vorgehensweise bei einer überwachten Klassifikation. Die Vorgehensweise wird für jeden Deskriptor und pro Deskriptor für alle möglichen Berechnungsparameter des Deskriptors (r_n und r_H für die geometrischen Deskriptoren) wiederholt.

Bei jeder Klassifikation werden Accuracy, Recall und Precision für die Klasse **Traube** berechnet. Im Falle der geometrischen Deskriptoren wird das jeweilige beste Ergebnis pro Deskriptor aus allen untersuchten Parameterwerten ermittelt. Das beste Ergebnis pro Deskriptor wird für den Vergleich aller Deskriptoren untereinander verwendet. Bei der Evaluierung pro geometrischen Deskriptor werden ausschließlich die fünf Ergebnisse mit der höchsten Accuracy betrachtet und aus ihnen das jeweils beste Ergebnisse ermittelt. Dabei gelten folgende Nebenbedingungen:

1. Einem hohen Recall wird Vorzug vor einer hohen Precision gegeben, da möglichst alle Trauben und Beeren detektiert werden sollen. Eine hohe Accuracy, die aus einem niedrigem Recall, aber einer sehr hohen Precision resultiert, wird daher verworfen.
2. Hohe Vollständigkeit (hoher Recall) bei einer hohen Fehlklassifikationsrate (geringe Precision) ist nicht zielführend, ebenso wie eine geringe Fehlklassifikationsrate bei geringer Vollständigkeit. Recall und Precision sollten daher ausgeglichen sein. Ist der Recall bereits hoch, wird ein etwas geringerer Recall im Vergleich zu einer anderen optimalen Radiankombination in Kauf genommen, falls dadurch eine merklich größere Precision erreicht wird.
3. Erzeugen mehrere Radiankombinationen ähnliche Ergebnisse, wird sich für die Kombination mit kleineren Radian entschieden, um Rechenzeit einzusparen.

Im Folgenden werden die Ergebnisse der Klassifikation für alle Deskriptoren vorgestellt. Zunächst erfolgt das Ergebnis für den HSV-Deskriptor. Im Anschluss erfolgt das Ergebnis für die SFH- und SFHc-Deskriptoren, welches beispielhaft erläutert wird, um die Vorgehensweise bei der Evaluierung der geometrischen Deskriptoren zu verdeutlichen. Die SF-, SFc- und CSHOT-Deskriptor werden darauffolgend in gleicher Weise evaluiert, die Darstellung der Ergebnisse erfolgt jedoch in verkürzter Form.

4.2.1.2 HSV-Deskriptor

Die Klassifikation der Testdaten mit dem HSV-Deskriptor erreicht eine Accuracy von 32%, mit einem Recall von 63,7% und einer Precision von 52,5%. Abbildung 4.12 zeigt das Klassifikationsergebnis. Nicht alle 3D-Punkte, die Trauben abbilden, wurden als Trauben klassifiziert. Punkte einer kleineren Traube auf rechten Seite wurden beinahe gänzlich fehlklassifiziert. Insbesondere Äste wurden fälschlicherweise als Traubepunkte klassifiziert und bilden wenige, aber größere zusammenhängende Regionen von Fehlklassifikationen. Dies ist vermutlich auf die stellenweise ähnlich braun-dunkle Färbung von Ästen und Beeren zurückzuführen. Blätter wurden hingegen weitgehend korrekt klassifiziert, Fehlklassifikationen von Blattpunkten sind punktuell oder äußern sich in nur kleinen Regionen von Fehlklassifikationen.



Abbildung 4.12: Initiale Klassifikation der Testdaten mit den HSV-Deskriptoren. Rot: Rest, Grün: Trauben.

4.2.1.3 SFH- und SFHc-Deskriptor

Tabelle 4.2 listet die fünf Radienkombinationen mit höchster Accuracy für den SFH-Deskriptor auf.

	Accuracy [%]	Recall [%]	Precision [%]
$r_n = 2\text{mm}; r_H = 12\text{mm}$	48,2	82,4	58,5
$r_n = 2\text{mm}; r_H = 17\text{mm}$	50,2	89,2	56,3
$r_n = 3\text{mm}; r_H = 12\text{mm}$	48,3	83,7	57,7
$r_n = 3\text{mm}; r_H = 17\text{mm}$	49,7	90,8	54,7
$r_n = 4\text{mm}; r_H = 17\text{mm}$	49,2	90,1	54,5

Tabelle 4.2: Ergebnisse der initialen Klassifikation für die SFH-Deskriptoren. Es werden lediglich die fünf Kandidaten mit höchster Accuracy dargestellt.

Eine dominante Accuracy kann nicht identifiziert werden, die Werte unterscheiden sich lediglich um maximal 2%-Punkte. Es zeigt sich jedoch, dass sich die Accuracy aus Recall und Precision Werten ergibt, die sich bis zu 8% unterscheiden können. Unter Anwendung der oben definierten Nebenbedingungen beträgt der optimale Nachbarschaftsradius r_n für die SFH-Deskriptoren 3mm , der optimale Histogrammradius $r_H = 17\text{mm}$. Es wird eine Accuracy von 49,7%, mit dem höchsten Recall von 90,8% und einer Precision von 54,7%, erhalten. Der höhere Recall dieser Radienkombination wird der etwas höheren Precision von 56,3% der Radienkombination $r_n = 2\text{mm}, r_H = 17\text{mm}$ vorgezogen.

Die SFH-Deskriptoren werden mit den HSV-Deskriptoren zum SFHc-Deskriptor konkateniert und erneut alle möglichen Radienkombinationen für die Klassifikation der Testdaten verwendet. Tabelle 4.3 listet die fünf Radienkombinationen mit höchster Accuracy für den SFHc-Deskriptor auf.

	Accuracy [%]	Recall [%]	Precision [%]
$r_n = 2mm; r_H = 27mm$	54,1	96,4	56,1
$r_n = 3mm; r_H = 12mm$	54,1	80,5	67,2
$r_n = 3mm; r_H = 17mm$	55,2	87,2	63,3
$r_n = 3mm; r_H = 22mm$	53,8	93,3	57,7
$r_n = 3mm; r_H = 27mm$	54,2	96,6	56,1

Tabelle 4.3: Ergebnisse der initialen Klassifikation für die SFHc-Deskriptoren. Es werden lediglich die fünf Kandidaten mit höchster Accuracy dargestellt.

Der optimale Nachbarschaftsradius der SFHc-Deskriptoren unter Berücksichtigung der Nebenbedingungen verbleibt identisch bei $r_n = 3mm$, der optimale Histogrammradius ebenfalls bei $r_H = 17mm$. Es wird eine maximale Accuracy von 55,2% erreicht, mit einem Recall von 87,2% und einer Precision von 63,3%. Die maximale Accuracy für den SFHc-Deskriptor ist damit um rund 5%-Punkte höher, als bei der ausschließlichen Verwendung des SFH-Deskriptors. Im Vergleich zur ausschließlichen Nutzung des SFH-Deskriptors erhöht sich die Precision um etwa 9%-Punkte, der Recall verringert sich jedoch geringfügig um etwa 3%-Punkte.



Abbildung 4.13: Initiale Klassifikation der Testdaten mit den SFHc-Deskriptoren. Rot: Rest, Grün: Trauben.

Der Recall ist für diese Radienkombination dennoch sehr hoch, bei einer deutlich gesteigerten Precision. Die anderen Radien besitzen ebenfalls hohe Werte für den Recall, jedoch sind Recall und Precision für diese Kombination am ausgeglichensten und der Histogrammradius r_H ist für diese Kombination am niedrigsten. Abbildung 4.13 zeigt das Klassifikationsergebnis. Fehlklassifikationen an den bräunlichen Ästen bei alleiniger Verwendung des HSV-Deskriptors (vgl. Abbildung 4.12) konnten durch die Verbindung von Geometrie und Farbe im SFHc-Deskriptor vermieden werden.

4.2.1.4 SF- und SFc-Deskriptor

Der optimale Nachbarschaftsradius r_n für die SF-Deskriptoren beträgt $8mm$, mit dem die Accuracy von 23,4% erhalten wird, bei einem Recall von 74,0% und einer Precision von 31,7%. Für die Untersuchung der Kombination des SF-Deskriptors mit dem HSV-Deskriptor zum SFc-Deskriptor $\mathbf{sfc} = [\mathbf{sf} \ \mathbf{hsv}]$ wird erneut mit allen Radien r_n klassifiziert. Der optimale Nachbarschaftsradius r_n erhöht sich in Kombination mit den HSV-Deskriptoren auf $10mm$. Die Accuracy erreicht 38,7%, mit einem höchsten Recall von 89,4% und einer Precision von 43,2%. Der Nachbarschaftsradius von $10mm$ dominiert bzgl. der Accuracy um 7 – 10%, was eine deutliche Steigerung gegenüber der Verwendung von ausschließlich SF-Deskriptoren darstellt.

4.2.1.5 CSHOT-Deskriptor

Der optimale Nachbarschaftsradius r_n für die CSHOT-Deskriptoren beträgt $3mm$, der optimale Histogrammradius $r_H = 17mm$. Die optimalen Radien decken sich somit mit den optimalen Radien für die SFH- und

SFHc-Deskriptoren. Offensichtlich liegen in diesem Radius Geometrien beider Klassen vor, die sich lokal am stärksten voneinander unterscheiden. Die dominante Accuracy erreicht 36,2%, mit dem höchsten Recall von 91,2% und einer Precision von 39,6%. Die Accuracy der anderen Radienkombinationen ist merklich um 3 – 4% geringer.

4.2.1.6 Zusammenfassung

Es wurden mehrere Deskriptoren hinsichtlich ihrer Unterscheidungskraft für die Klassifikation von Weinreben mittels der Metriken Accuracy, Recall und Precision untersucht. Tabelle 4.4 listet die besten Ergebnisse pro Deskriptor vergleichend auf. Die SFHc-Deskriptoren erweisen sich demnach als Deskriptor, mit dem das beste Klassifikationsergebnis erzeugt werden kann. Kein anderer Deskriptor erreichte ähnliche hohe Werte für Accuracy, Recall und Precision. Recall und Precision bei Verwendung der SFHc-Deskriptoren sind im Vergleich zu den anderen Deskriptoren ausgeglichen und hoch. Mit Ausnahme der SFH-Deskriptoren erreichen alle anderen Deskriptoren deutlich geringere Werte in Recall und Precision. Werden ausschließlich SFH-Deskriptoren verwendet, wird die zweithöchste Accuracy von 49,7% erreicht. In Kombination mit den HSV-Deskriptoren zum SFHc-Deskriptor erhöht sich die Accuracy der Klassifikation merklich um mehr als 5%-Punkte auf 55,2%. Dieser vorteilhafte Effekt ist kongruent zu den Ergebnissen der Klassifikation mit den SF-Deskriptoren. Es lässt sich daher zusätzlich konstatieren, dass die Kombination von geometrischen und farblichen Deskriptoren einen robusteren Deskriptor für die Klassifikation von Weinreben erzeugt.

	Accuracy [%]	Recall [%]	Precision [%]
<i>HSV</i>	32,0	63,7	52,2
<i>SFH</i>	49,7	90,8	54,7
SFHc	55,2	87,2	63,3
<i>SF</i>	23,4	74,0	31,7
<i>SFc</i>	38,7	89,4	43,2
<i>CSHOT</i>	36,2	91,2	39,6

Tabelle 4.4: Beste Ergebnisse der initialen Klassifikation für alle Deskriptoren.

Die mit dem SFHc-Deskriptor erreichten Werte für Recall und Precision erlauben die Detektion des Großteils der Punkte der vorhandenen Trauben bei einer möglichst geringen Anzahl von Fehlklassifikationen. Die SFHc-Deskriptoren werden daher für alle weiteren Untersuchungen dieser Arbeit als Deskriptor für die Klassifikation von Weinreben verwendet. Dieses Ergebnis deckt sich mit den Ergebnissen von Mack et al. [2017b], welche die FPFHs von Rusu [2009] als besten geometrischen Deskriptor zur Klassifizierung von Punktwolken von Weinreben und anderen Pflanzen im Vergleich mit anderen geometrischen Deskriptoren wie bspw. SHOT identifizieren konnten.

4.2.2 Weiterführende Untersuchungen

Nach der Identifikation des SFHc-Deskriptors als besten Deskriptor für die Klassifikation der Weinreben werden weitere Untersuchungen hinsichtlich der optimalen Radienkombination, des Einflusses der Klassenglättung auf das Klassifikationsergebnis und die Bestimmung einer optimalen Radienkombination für das BBCH75 Reifestadium durchgeführt.

4.2.2.1 Einfluss von GCO auf die Bestimmung optimaler Radian

Im vorherigen Unterkapitel wurden der SFHc-Deskriptor als derjenige Deskriptor identifiziert, der ohne weitere Klassifikationsmaßnahmen wie GCO die größte Unterscheidungskraft besitzt. Um ein optimale Klassifikationsergebnis des SFHc-Deskriptors hinsichtlich Recall und Precision zu erreichen, wurden vorab optimale

Radien ermittelt. Diese optimale Radienkombination erzeugt bei ausschließlicher Verwendung der IVM als Klassifikator das beste Ergebnis im Sinne einer binären Klassenzuweisung nach der jeweils höchsten Wahrscheinlichkeit aus $W_p = [W_1, W_2]$.

GCO nutzt die Wahrscheinlichkeiten der Klassenzugehörigkeit $W_p = [W_1, W_2]$ und die zugewiesene Klasse c_p jedes 3D-Punktes zur Minimierung der für die Klassifikation nötigen Energie (s. Kapitel 3.2.6 GCO). Die Klassifikation mit der IVM war zwar optimal hinsichtlich der binären Klassenzuweisung, die durch W_p und c_p entstandene Energie muss jedoch nicht zwangsläufig optimal für die Klassenglättung mit GCO sein.

In diesem Kapitel soll daher der Einfluss der Klassenglättung mit GCO auf das Klassifikationsergebnis unter Einsatz mehrerer Radienkombinationen für den SFHc-Deskriptor von r_n und r_H erneut untersucht werden. Ziel dieses Unterkapitels ist die Untersuchung des Einflusses von GCO auf das Klassifikationsergebnis und die Bestimmung einer optimalen Radienkombination hinsichtlich Recall und Precision bei Einsatz von GCO.

Für r_n werden wie bisher Werte im Bereich $[1mm..5mm]$ mit $1mm$ verwendet. In den vorherigen Untersuchungen wurden die besten Ergebnisse des SFHc-Deskriptors mit dem Histogrammradius $r_H = 17mm$ erreicht. Jedoch erzeugte auch ein Histogrammradius $r_H = 12mm$ bereits eine hohe Accuracy. Für r_H werden daher Werte im Bereich $[2..17mm]$ mit $1mm$ Schrittweite verwendet, anstelle wie vormals in $5mm$ Schritten. Durch die Reduzierung des Wertebereichs auf maximal $17mm$ und der Schrittweite auf $1mm$ soll zusätzlich festgestellt werden, ob bisher nicht betrachtete Histogrammradien $r_H < 17mm$ zu einer ebenso hohen oder sogar zu einer Steigerung der Accuracy führen können.

Für GCO wird ein Nachbarschaftsradius $r_G = 5mm$ um einen Quellpunkt $\mathbf{p} \in P^t$ für den Glattheitsterm gewählt. Als Ähnlichkeitsmaß w_{pq} zwischen Nachbarpunkten $\mathbf{q} \in P^k$ im Radius r_G um den Quellpunkt wird Gleichgewichtung gewählt, d. h. $w_{pq} = 1$, wenn Punkte benachbart sind. Die Größe des Nachbarschaftsradius r_G wurde entsprechend der räumlichen Ausdehnung falsch klassifizierter Punkte auf Trauben ausgewählt. Die Klassenglättung soll Punkten auf Trauben, die fälschlicherweise als **Rest** klassifizierten wurden, der Klasse **Traube** zuweisen und umgekehrt auf allen anderen Objekten als **Traube** klassifizierte Punkte der Klasse **Rest**. Ein zu kleiner betrachteter Bereich enthielte u. U. primär identisch klassifizierte Punkte. Ohne eine Hinzunahme von Punkten der richtigen Klassen existiert jedoch kein Anhaltspunkt für die Korrektur der falschen Klasse und würde zu keiner Verbesserung der Klassifikation führen. Bei einem zu groß gewählten Radius könnten hingegen die richtig klassifizierten Beeren, die nur einen geringen Anteil der Punkte der gesamten Punktwolke darstellen, als Abweichung von der zahlenmäßig größeren Klasse **Rest** interpretiert und in der Folge falsch klassifiziert werden. Problematisch sind dabei vorallem Punkte von Trauben, die in der Nähe von Blättern oder Ästen liegen. Ein Radius von $r_G = 5mm$ entspricht ungefähr dem Ausmaß einer Beere und enthält genügend korrekt klassifizierte Punkte beider Klassen, um falsch klassifizierte Punkte zu korrigieren.

Für die Untersuchungen werden dieselben Testdaten und Trainingsdaten wie zuvor verwendet. Die Vorgehensweise bei der Bestimmung der optimalen Radienkombination ist ebenfalls identisch, die Klassifikation wird nun lediglich inklusive der Klassenglättung mit GCO durchgeführt. Tabelle 4.5 zeigt die besten fünf Ergebnisse nach der Klassenglättung.

	Accuracy [%]	Recall [%]	Precision [%]
$r_n = 3mm; r_H = 9mm$	70,7	89,0	79,4
$r_n = 3mm; r_H = 11mm$	70,7	85,6	82,6
$r_n = 3mm; r_H = 12mm$	73,7	85,2	86,4
$r_n = 3mm; r_H = 14mm$	71,4	94,5	75,8
$r_n = 4mm; r_H = 16mm$	72,5	95,8	75,7

Tabelle 4.5: Ergebnisse für die Klassifikation mit dem SFHc-Deskriptor nach der Klassenglättung mit GCO. Es werden lediglich die fünf Kandidaten mit höchster Accuracy dargestellt.

Ohne Klassenglättung betrug die höchste Accuracy 55,2% für die SFHc-Deskriptoren. Die Klassenglättung mit GCO erhöht die Werte der Accuracy deutlich um ca. 15 – 18%-Punkte auf bis zu 73,7%. Insbesondere wird die Precision von vormals 63,3% um mind. 12%-Punkte erhöht. Der Recall erfährt teilweise Erhöhungen

um bis zu 8%-Punkte auf 95,8% (vormals 87,2%), sinkt jedoch auch teilweise um bis zu 2%-Punkte auf 85,2%. Die Accuracy-Werte unterscheiden sich untereinander um maximal 3%-Punkte.

Im Vergleich zu den vorherigen Untersuchungen werden die höchsten Accuracy-Werte ebenfalls größtenteils in einem Nachbarschaftsradius r_n von 3mm erreicht. Der optimale Histogrammradius r_H sinkt jedoch im Vergleich von 13 – 17mm auf Werte zwischen 9mm und 16mm. Die Ergebnisse zeigen, dass bei Verwendung von GCO die Ermittlung einer optimalen Radienkombination für eine gegebene Aufgabe stets mit der Klassenglättung durchgeführt werden sollte.

Die höchste Accuracy wird bei $r_n = 3mm$ und $r_H = 12mm$ erzeugt, jedoch liegt hier der Recall niedriger als die Precision. Das ausgeglichene Verhältnis von Recall und Precision wird bei $r_n = 3mm$ und $r_H = 9mm$ erreicht. Der Recall liegt hier mit 89% sehr hoch. Die Precision liegt etwas niedriger als der Recall mit knapp 80%, ist jedoch höher als für Histogrammradien $> 12mm$. Die Radienkombination $r_n = 3mm$ und $r_H = 9mm$ wird unter Berücksichtigung der oben genannten Bedingungen daher als optimale Radienkombination für alle künftigen Klassifikationen von Weinreben im BBCH89 Reifestadium verwendet. Abbildung 4.14 zeigt die Testdaten, die initiale Klassifikation mit $r_n = 3mm$ und $r_H = 9mm$ und das Ergebnis für $r_n = 3mm$ und $r_H = 9mm$ nach der Klassenglättung. Falsch klassifizierte Punkte werden durch GCO korrigiert, die Klassenregionen im Vergleich zur initialen Klassifikation homogener. Nur wenige, kleinere Regionen mit falsch klassifizierten Punkten verbleiben.

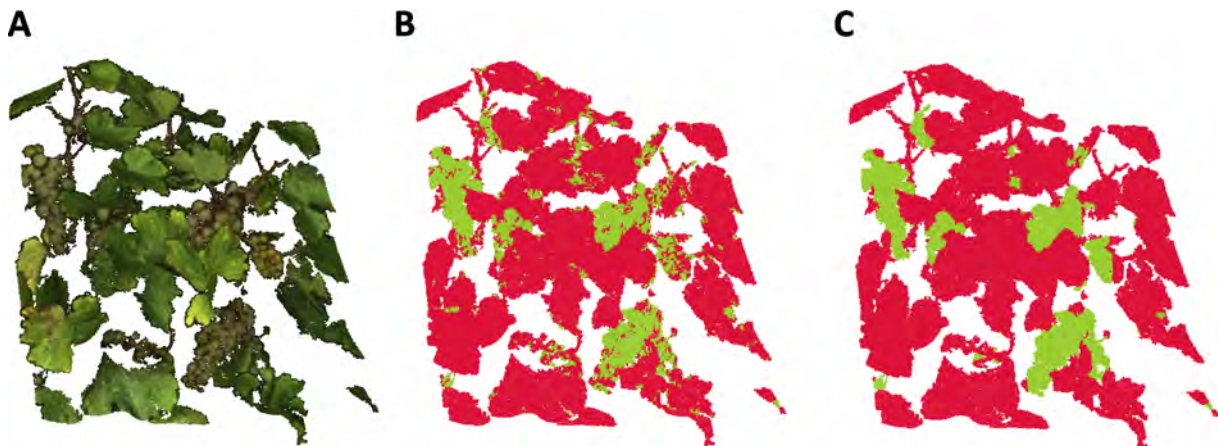


Abbildung 4.14: A: Testdaten in RGB-Farben. B: Initiale Klassifikation der Testdaten mit den SFHc-Deskriptoren. C: Klassifikationsergebnis nach der Klassenglättung. Rot: Rest, Grün: Trauben.

Da Geometrie und Farbe der Weinreben für beide Erziehungssysteme in diesem kleinen lokalen Bereich von maximal 9mm Radius sehr ähnlich sind, wird diese Radienkombination unabhängig vom Erziehungssystem verwendet. Durch den geringen Histogrammradius von 9mm wird des Weiteren Berechnungszeit eingespart.

4.2.2.2 Ermittlung einer optimalen Radienkombination für BBCH75

Die bisherigen Untersuchungen fanden auf Daten der MSS-Anbaureihe im Reifestadium BBCH89 statt. Da sich Geometrie und Farbe von Blattwerk und Trauben zwischen dem BBCH75 und BBCH89 Reifestadium deutlich unterscheiden, muss für das BBCH75 Reifestadium erneut eine optimale Radienkombination für die SFHc-Deskriptoren ermittelt werden. Dabei wird sich auf die Erkenntnisse aus den vorherigen Untersuchungen bezogen. Aus diesen wurde Folgendes ersichtlich:

1. Der optimale Histogrammradius r_H sollte im Wertebereich bis maximal 17mm gesucht werden.
2. Die Klassenglättung mit GCO verbessert die Accuracy deutlich um bis zu 18%-Punkte.

3. Die Suche nach optimalen Radien sollte inklusive der Klassenglättung durchgeführt werden.

Um die geringere Größe der Beeren in diesem Reifestadium zu berücksichtigen, werden für r_n Werte im Bereich $[1mm..3mm]$ mit $1mm$ Schrittweiten und für r_H Werte im Bereich $[2..15mm]$ mit $1mm$ Schrittweiten verwendet. Aus einer vollständig vorverarbeiteten Punktwolke einer TS-Anbaureihe im BBCH75 Reifestadium werden Testdaten $\mathcal{P}^t$ und Trainingsdaten $T \notin \mathcal{P}^t$ ausgewählt und für $\mathcal{P}^t$ und T jeweils die SFHc-Deskriptoren aller möglichen Radienkombinationen berechnet. Die Testdaten werden anschließend durch die IVM und GCO pro Radienkombination klassifiziert.

Die Ergebnisse werden wie in den vorherigen Untersuchungen bewertet. Lediglich die fünf Radienkombinationen mit der höchsten Accuracy werden unter Berücksichtigung der Nebenbedingungen bewertet und aus ihnen eine optimale Radienkombination ermittelt. Tabelle 4.6 listet die fünf Radienkombinationen mit höchster Accuracy auf.

	Accuracy [%]	Recall [%]	Precision [%]
$r_n = 2mm; r_H = 11mm$	75,7	91,9	82,4
$r_n = 2mm; r_H = 14mm$	77,1	94,5	81,5
$r_n = 3mm; r_H = 12mm$	74,8	97,5	73,3
$r_n = 3mm; r_H = 13mm$	76,0	95,2	79,8
$r_n = 3mm; r_H = 14mm$	75,7	97,4	77,7

Tabelle 4.6: Ergebnisse nach der Klassenglättung für die Testdaten im BBCH75 Reifestadium. Es werden lediglich die fünf Kandidaten mit höchster Accuracy dargestellt.

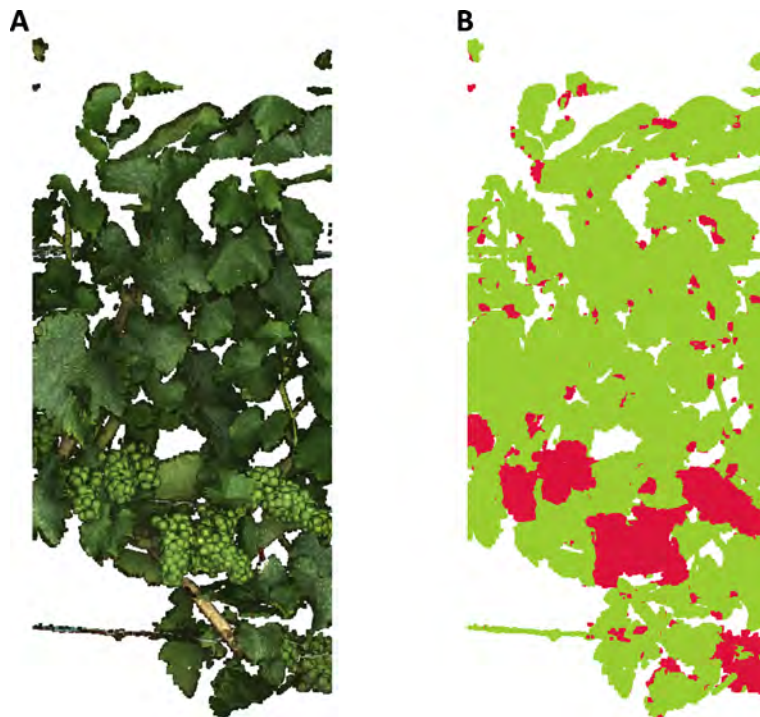


Abbildung 4.15: **A:** Testdaten in RGB-Farben. Die Maße des Ausschnitts betragen rund $40 \times 100 \times 60 cm$. Trauben und Blätter sind sich farblich sehr ähnlich und in einem durchgehend hellgrünen Farbton. **B:** Das Klassifikationsergebnis mit SFHc-Deskriptoren für $r_n = 2mm$, $r_H = 14mm$. Rot: Trauben; Grün: Rest

Abbildung 4.15 zeigt das Klassifikationsergebnis bei Verwendung der optimalen Radienkombination. Alle Ergebnisse besitzen einen Recall $> 90\%$ mit hohen Werten zwischen $73 - 82\%$ für die Precision. Ab einem Nachbarschaftsradius $r_n = 3mm$ wächst der Recall, die Precision sinkt jedoch unter 80% . Die höchste

Accuracy beträgt 77,1% bei ausgeglichenen Werten für Recall mit 94,5% und einer Precision von 81,5%. Dieses Ergebnis wird für einen Nachbarschaftsradius $r_n = 2mm$ und einen Histogrammradius $r_H = 14mm$ erreicht. Da Geometrie und Farbe in beiden Erziehungssystemen innerhalb dieser geringen Radien lokal sehr ähnlich sind, wird diese Radienkombination künftig für beide Erziehungssysteme in diesem Reifestadium verwendet.

4.2.2.3 Zusammenfassung

Es konnte demonstriert werden, dass die Klassenglättung mit GCO die Accuracy um mehrere Prozentpunkte verbessert. Die ermittelten optimalen Radien r_n und r_H des SFHc-Deskriptors unterschieden sich von den optimalen Radien aus den Untersuchungen zur Identifikation des besten Deskriptors, bei denen noch keine Klassenglättung vorgenommen wurde. Die Suche nach optimalen Radien sollte daher stets *inklusive* der Klassenglättung durchgeführt werden.

Es wurde eine optimale Radienkombination für die Klassifikation von Weinreben im BBCH75 Reifestadium unter Berücksichtigung der eben beschriebenen Erkenntnisse ermittelt. Es zeigte sich, dass für das BBCH75 und BBCH89 Reifestadium aufgrund ihrer unterschiedlichen Morphologie unterschiedliche Radienkombinationen verwendet werden müssen.

4.2.3 Klassifikation von 5m Ausschnitten beider Erziehungssysteme und Reifestadien

Im Folgenden werden 5m lange Ausschnitte der rekonstruierten 25m langen Anbaureihen unter Nutzung des SFHc-Deskriptors mit der vorher bestimmten optimalen Radienkombination für beide Reifestadien klassifiziert. Um Recall und Precision berechnen und einen Eindruck von der Klassifikationsgenauigkeit bei längeren Ausschnitten gewinnen zu können, werden für jeweils 2,5m der 5m-Ausschnitte zusätzlich die Klasse jedes 3D-Punktes manuell bestimmt. Ein Ausschnitt von 2,5m Länge besitzt nach allen Vorverarbeitungsschritten noch ca. zu klassifizierende 0,8 – 1,5 Millionen Punkte. Zunächst werden die Anbaureihen im Reifestadium BBCH89 betrachtet, die Anbaureihen im Reifestadium BBCH75 folgen darauf.

4.2.3.1 Anbaureihen im Reifestadium BBCH89



Abbildung 4.16: Anbaureihe der TS-Erziehung von ungefähr 2,5m Länge, aus Rose et al. [2016].

Abbildung 4.16 zeigt einen Ausschnitt aus der Anbaureihe von ungefähr 2,5m Länge nach der TS-Erziehung. Tabelle 4.7 fasst die Ergebnisse der Klassifikation der Anbaureihen beider Erziehungssysteme zusammen.

	Recall [%]	Precision [%]
TS89 initiale Klassifikation:	87	39
TS89 nach Klassenglättung:	94	63
MSS89 initiale Klassifikation:	81	23
MSS89 nach Klassenglättung:	89	39

Tabelle 4.7: Ergebnisse der Klassifikation für die Ausschnitte der TS- und MSS-Erziehung im BBCH89 Reifestadium.

Für die initiale Klassifikation der TS-Anbaureihe wird ein Recall von 87% und eine Precision von 39% erreicht. Abbildung 4.17 zeigt den klassifizierten Ausschnitt nach der Klassenglättung. Die Klassenglättung mit GCO erhöht den Recall auf 94% und die Precision auf 63%. Klassifikationsrauschen ist nur in geringem Ausmaß erkennbar.

Die Werte für Recall und Precision der TS-Anbaureihe sind bzgl. des Recalls mit 94% ähnlich zu den Werten der Klassifikation der Testdaten im Kapitel 4.2.1 der Evaluation der Deskriptoren (Recall: 89,0%; Precision: 79,4%), die Precision ist jedoch um rund 17%–Punkte niedriger. Dies ist höchstwahrscheinlich auf die im Vergleich zu den Testdaten größere Länge der betrachteten Ausschnitte zurückzuführen, wodurch sich der Anteil von Punkten der Klasse **Rest** im Verhältnis zur Klasse **Traube** vergrößert und eine höhere Rate von Fehlklassifikationen für die Klasse **Traube** wahrscheinlicher wird.

Abbildung 4.18 zeigt einen Ausschnitt aus der Anbaureihe nach MSS-Erziehung im Reifestadium BBCH89. Bei der initialen Klassifikation der MSS-Anbaureihe wird ein Recall von 81% und eine Precision von 23% erreicht. Nach der Klassenglättung erhöht sich der Recall auf 89% und die Precision auf 39%. Der Recall ist hoch, jedoch liegt die Precision im Vergleich zur in Kapitel 4.2.1 erreichten Precision mit 39% um 40%-Punkte niedriger, obwohl die Testdaten dort ebenfalls aus der MSS-Anbaureihe entnommen wurden. Dieser Umstand resultiert 1. aus dem größeren betrachteten Ausschnitt und 2. aus dem für MSS-Anbaureihen typischen, sehr dichten Blattwerk. Objekte liegen hier sehr nah zusammen und überlappen sich. Aufgrund des hohen Verdeckungsgrades können Objekte zudem in weniger Bildern beobachtet werden. Dies resultiert in nur teilweise rekonstruierten Objekten und in teilweise korumpierten Oberflächengeometrien, welche für die Klasse untypische SFHc-Deskriptorvektoren erzeugen. Abbildung 4.19 zeigt den gleichen Ausschnitt der MSS Anbaureihe nach der Klassenglättung. Die Regionen einzelner Klassen bilden homogene Regionen, das Klassifikationsrauschen ist gering. Stützpfeiler und Drähte wurden korrekt klassifiziert. Es existieren jedoch kleinere Regionen von Fehlklassifikationen, welche die Precision verringern.



Abbildung 4.17: Punktwolke der TS-Erziehung nach der Klassenglättung (Rot: **Rest**, Grün: **Trauben**), aus Rose et al. [2016].



Abbildung 4.18: Ausschnitt aus der Anbaureihe der MSS-Erziehung.



Abbildung 4.19: Der Ausschnitt der MSS-Anbaureihe nach der Klassenglättung (Rot: Rest, Grün: Trauben).

4.2.3.2 Anbaureihen im Reifestadium BBCH75

	Recall [%]	Precision [%]
TS75 initiale Klassifikation:	92	43
TS75 nach Klassenglättung:	96	59
MSS75 initiale Klassifikation:	93	27
MSS75 nach Klassenglättung:	99	50

Tabelle 4.8: Ergebnisse der Klassifikation für den Ausschnitt nach der TS- und MSS-Erziehung im BBCH75 Reifestadium.

Tabelle 4.8 fasst die Ergebnisse der Klassifikation der Anbaureihen nach TS- und MSS-Erziehung im BBCH75 Reifestadium zusammen. Im Reifestadium BBCH75 wird für die TS-Anbaureihe ein hoher initialer Recall von 92% und eine Precision mit 43% erreicht. Die Klassenglättung verbessert Recall und Precision der initialen Klassifikation um mehrere Prozentpunkte, insbesondere die Precision verbessert sich um 16%-Punkte auf 59%. Die Werte sind denen der Klassifikation der TS-Anbaureihe im Reifestadium BBCH89 ähnlich (Recall: 94%, Precision 63%).

Im Vergleich mit den Testdaten im Reifestadium BBCH75 aus Kapitel 4.2.1 (Recall: 94,5%, Precision 81,5%) ist primär die Precision deutlich niedriger. Wie in der Klassifikation der Ausschnitte im BBC89 Reifestadium resultiert dieses Phänomen höchstwahrscheinlich aus dem vergleichsweise größeren betrachteten Ausschnitt, wodurch Fehlklassifikationen wahrscheinlicher werden.

Abbildung 4.20 visualisiert den Ausschnitt einer TS-Anbaureihe im BBCH75 Reifestadium von ungefähr 2,5m Länge. Neben den Weinreben sind Stützpfeosten, Drähte und Markierungsmaterial erkennbar.



Abbildung 4.20: Ausschnitt der TS-Anbaureihe im Reifestadium BBCH75.

Abbildung 4.21 zeigt den klassifzierten Ausschnitt nach der Klassenglättung. Der Großteil der Punkte wurde korrekt klassifiziert, kleinere Regionen von Fehlklassifikationen verbleiben jedoch.



Abbildung 4.21: Ausschnitt der TS-Anbaureihe im BBCH75 Reifestadium nach der Klassenglättung (Rot: Trauben, Grün: Rest).

Die initiale Klassifikation der MSS-Anbaureihe erreicht einen Recall von 93% und eine Precision von 27%. Die Klassenglättung verbessert die initiale Klassifikation deutlich. Insbesondere der Recall ist nach der Klassenglättung mit 99% sehr hoch. Damit wurden nahezu alle Punkte, die Trauben abbilden, anhand der Klassifikation detektiert. Die Precision steigt auf 50%. Insgesamt ist die Precision niedriger als die Precision bei der TS-Anbaureihe in beiden Reifestadien.

Der Unterschied in der Precision zwischen beiden Erziehungssystemen lässt sich wie schon im BBCH89 Reifestadium aus dem dichteren Blattwerk des MSS-Erziehungssystems erklären, welches zur häufigeren Überlagerung von Objekten und einer dadurch korumpierten lokalen Geometrie führt, welche die Klassifikation erschwert. Im Vergleich mit der MSS-Anbaureihe im BBCH89 Reifestadium (Recall: 89%, Precision 39%) sind Recall und Precision im BBCH75 Reifestadium jedoch merklich höher. Ein möglicher Grund dafür könnte sein, dass das Blattwerk in diesem Reifestadium noch nicht so dicht liegt, wie in späteren Reifestadien.

Abbildung 4.22 zeigt den Ausschnitt der MSS-Anbaureihe im Reifestadium BBCH75. Einige Objekte vor allem im unteren Teil der Anbaureihe wurden nur teilweise rekonstruiert. Die erhöhte Dichte des Blattwerks bei der MSS-Erziehung erschwert auch in diesem Reifestadium die Beobachtung von Objekten durch einen größeren Verdeckungsgrad.



Abbildung 4.22: Ausschnitt der MSS-Anbaureihe im Reifestadium BBCH75.

Abbildung 4.23 zeigt das Klassifikationsergebnis nach der Klassenglättung. Kleinere Regionen von Fehlklassifikationen verbleiben, Drähte und Markierungsmaterial wurden korrekt klassifiziert.



Abbildung 4.23: Ausschnitt nach der Klassenglättung (Rot: Trauben, Grün: Rest).

4.2.3.3 Zusammenfassung

Bei beiden Erziehungssystemen und beiden Reifestadien werden hohe Werte für den Recall zwischen 89–98% erreicht. Bei keiner einzigen rekonstruierten Traube wurde alle 3D-Punkte falsch klassifiziert. Stattdessen wurden nur wenige 3D-Punkte einer rekonstruierten Traube nicht korrekt klassifiziert, die zusammen jedoch niemals eine ganze Traube konstituieren. Es lässt sich daher konstatieren, dass alle rekonstruierten Trauben in der Punktwolke durch die Klassifikationsstrategie detektiert wurden und für die weiteren Schritte der Auswertung genutzt werden können.

Die Precision nach der Klassenglättung liegt bei 38 – 62% und ist stets kleiner als der Recall für beide Erziehungssysteme und Reifestadien. Große Unterschiede zeigen sich zwischen beiden Erziehungssystemen. Die Precision ist für das TS-System stets höher als für das MSS-System. Dies ist höchstwahrscheinlich auf die höhere Dichte von Blättern, Trauben und Ästen in den MSS-Anbaureihen zurückzuführen. Objekte überlagern sich dort und korrumpieren die Geometrie der Objekte. Kleinere Regionen von Fehlklassifikationen verbleiben auch nach der Klassenglättung.

Es konnte gezeigt werden, dass die Klassenglättung mit GCO die Ergebnisse der initialen Klassifikation auch bei der Klassifikation größerer Punktwolken deutlich steigert. Klassifikationsrauschen wird durch GCO verringert und homogene Klassenregionen erhalten. Dies ist besonders für die spätere Beerendetektion wichtig, da dafür möglichst alle 3D-Punkte einer Beere zur Verfügung stehen sollten.

4.3 Quantifikation

Im zurückliegenden Kapitel wurden für beide Erziehungssysteme und Reifestadien jeweils 5m Ausschnitte vollständig klassifiziert. In diesem Kapitel wird die Anzahl der Trauben und Beeren, sowie der Beerengröße in diesen Ausschnitten bestimmt. Um diese Größen zu bestimmen, werden ausschließlich diejenigen 3D-Punkte der vier Ausschnitte genutzt, die der Klasse **Traube** zugewiesen wurden. Pro Ausschnitt sei diese Punktwolke jeweils als $\mathcal{P}^T$ bezeichnet.

In einem ersten Schritt werden die Punkte $\mathcal{P}^T$ mittels eines *Connected Components*-Verfahrens (CC) in N-Punktregionen segmentiert, die zunächst als Traubenkandidaten gelten. Diese fungieren als Basis für die weitere Quantifikation der Ertragsparameter. Während des CC-Verfahrens werden kleinere Regionen gelöscht, die heuristisch als zusammenhängende, fehlklassifizierte Punkte interpretiert werden. In einem zweiten Schritt werden die ermittelten Traubenkandidaten einzeln betrachtet. In ihnen wird mittels des Programms *Berryfinder* über Kugelapproximation nach Beeren gesucht und dabei anhand der Anzahl detektierter Beeren überprüft, ob die Traubenkandidaten tatsächlichen Trauben entsprechen. Traubenkandidaten werden entweder als Traube verworfen oder akzeptiert. Die Anzahl der in den akzeptierten Traubenkandidaten detektierten Beeren gelten als Beerenanzahl, die Anzahl akzeptierter Traubenkandidaten als Anzahl der Trauben im betrachteten Ausschnitt. Der Radius der Kugeln gilt als Beerengröße.

Zunächst wird das Connected Components-Verfahren beschrieben. Anschließend erfolgt die Beschreibung des Algorithmus' *Berryfinder* zur Detektion von Beeren in den Traubenkandidaten und eine Evaluierung des Berryfinders anhand von Referenzdaten. Im Anschluss erfolgt die detaillierte Schilderung der Ergebnisse der Detektion der sichtbaren Trauben und Beeren, sowie der Schätzung der Beerengröße für die vier Anbaureihen und eine Evaluierung der Ergebnisse anhand von manuell erhobenen Referenzwerten.

4.3.1 Detektion sichtbarer Beeren und Trauben

4.3.1.1 Identifikation von Traubenkandidaten mit Connected Components

Punkte in $\mathcal{P}^T$ sind bisher noch ungeordnete Punktregionen und geben für sich noch keinerlei Auskunft darüber, wie groß die Anzahl der Trauben oder Beeren im betrachteten Ausschnitt ist, lediglich, dass ein 3D-Punkt zur Klasse **Traube** gehört. Es kann jedoch angenommen werden, dass dicht zusammenliegende

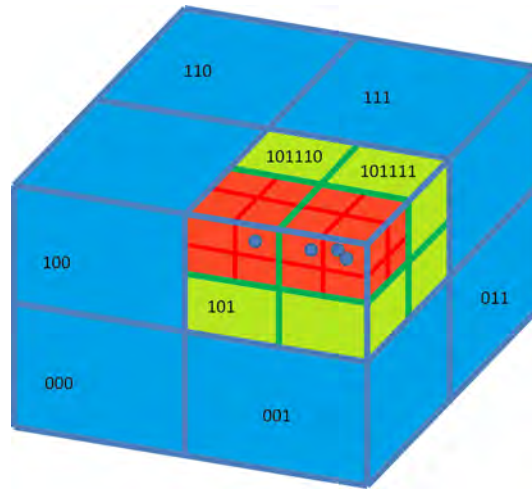


Abbildung 4.24: Der Wurzelraum (blau) wird in acht Oktanten (000 - 111) zentrisch zu seinem Schwerpunkt aufgeteilt (010 nicht sichtbar). Enthält ein Oktant 3D-Punkte (blaue Kreise), wird dieser *Vorfahr-Oktant* erneut in acht *Kinder-Oktanten* untergliedert (grün), die sich wiederum zentrisch um den Mittelpunkt des Vorfahr-Oktanten anordnen (101 000 - 101 000). Die Untergliederung geschieht rekursiv für jeden Oktanten, bis ein Oktant entweder keinen 3D-Punkt mehr enthält oder die maximale Tiefe D (hier ist $D = 3$) erreicht wurde (rot).

Punkte in $\mathcal{P}^T$ mit hoher Wahrscheinlichkeit Trauben entsprechen. Es sollen daher in $\mathcal{P}^T$ Punktregionen identifiziert werden, deren Punkte potentiell eine einzelne Traube abbilden. Des Weiteren verbleiben auch nach der Klassenglättung mit GCO weitere fehlklassifizierte Punktregionen, die tatsächlich der Klasse **Rest** angehören. Diese Fehlklassifikationen sollen aus $\mathcal{P}^T$ entfernt werden.

Das Connected Components Verfahren (CC-Verfahren) ermöglicht die Detektion dicht zusammen liegender Punkte und fasst sie semantisch als ein einziges Objekt zusammen. Der in dieser Arbeit genutzte Algorithmus ist in CloudCompare integriert und basiert auf einem sog. *Octree*. Ein Octree beginnt mit einem ursprünglichen, würfelförmigen 3D-Raum der Seitenlänge L , der als Wurzelraum bezeichnet wird. Der Wurzelraum ist dabei parallel zu den Achsen des Koordinatensystems von $\mathcal{P}^T$ und punktsymmetrisch um den Schwerpunkt von $\mathcal{P}^T$ angeordnet. Ein Octree untergliedert den Wurzelraum in rekursiv kleiner werdende, ebenfalls würfelförmige *Oktanten*, die sich entweder Kanten, Ecken oder Flächen teilen, sich jedoch nicht überlappen.

Der Wurzelraum wird zunächst in acht Oktanten mit der Seitenlänge $L/2$ unterteilt. Der untergliederte Raum wird als *Vorfahr*, die neu entstehenden acht Oktanten entsprechend als *Kinder* bezeichnet. Die Kinder-Oktanten sind dabei zentrisch um den Schwerpunkt von $\mathcal{P}^T$ angeordnet und berühren ihn mit einer Ecke. Jeder Kinder-Oktant des Wurzelraums kann nun wiederum rekursiv in weitere Oktanten untergliedert werden. Die Anzahl möglicher Untergliederungen eines Wurzelraumes in kleiner werdende Oktanten wird als Tiefe D bezeichnet und vom Nutzer definiert. Die kleinste mögliche Seitenlänge L_{min} eines Oktanten für einen Octree der Tiefe D ergibt sich über $L_{min} = L/2^D$.

Ein Oktant kann eine beliebige Anzahl von Punkten umschließen. Enthält ein Oktant mind. einen 3D-Punkt bezeichnet man ihn als *besetzt*. Ist ein Oktant besetzt, wird er erneut in acht Oktanten untergliedert, die zentrisch zum Mittelpunkt des Oktanten angeordnet sind. Die Untergliederung geschieht für alle besetzten Oktanten solange die maximale Tiefe D noch nicht erreicht wurde. Leere Oktanten werden nicht erneut untergliedert. Die maximale Anzahl möglicher Oktanten mit mindestens einem enthaltenem 3D-Punkt ist 8^D .

Die Position der acht Oktanten wird dabei über einen *3bit-Positions Wert* (000 - 111) relativ zum Schwerpunkt seines Vorfahr-Oktanten angegeben. Abbildung 4.24 zeigt die Anordnung der blauen Oktanten 000 - 111 des Wurzelraumes. Die Positionsvergabe folgt für jede weitere Untergliederung dieser Anordnung. Die Position eines Oktanten in der Tiefe D wird durch Konkatenation des 3bit-Werts des Vorfahren und des eigenen 3bit-Werts angegeben (beispielhaft: grüne Oktanten 101-110 und 101-111 der Tiefe $D = 2$).

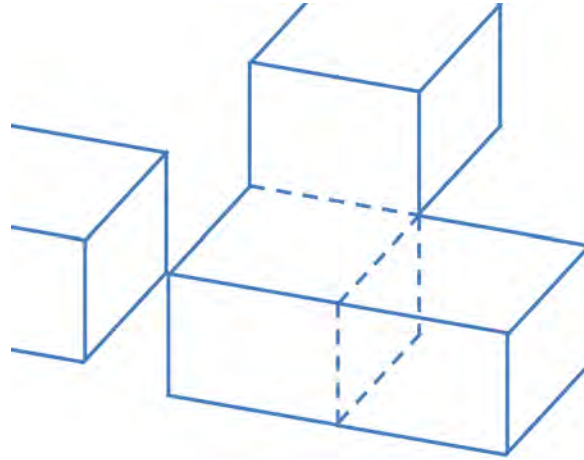


Abbildung 4.25: Die möglichen Nachbarschaftsrelationen eines Oktanten. Es existieren acht Nachbarschaften über geteilte Ecken, sechs Nachbarschaften über geteilte Flächen und zwölf Nachbarschaften über geteilte Kanten.

Die Nachbarn eines Oktanten werden bspw. nach den Methoden in Lumia et al. [1983] mittels Look-Up-Tabellen (LUT) bestimmt. Jeder Oktant kann dabei eindeutig über seinen 3bit-Positionswert identifiziert werden. Der Wurzelraum wird bis zur Tiefe D in Oktanten der Seitenlänge L_{min} untergliedert. Zwei besetzte Oktanten gelten als *direkt benachbart*, wenn sie entweder eine Seitenfläche, eine Kante oder eine Ecke teilen (Abbildung 4.25). Ein Oktant besitzt demnach bis zu 26 direkte Nachbarn.

Alle besetzten direkten K-Nachbaroktanten eines Oktanten U seien in der Menge $Okt^k = \{Okt_0, \dots, Okt_k\}$ $k = 1, \dots, K$ zusammengefasst. Sie gelten unmittelbar als *verbundene Komponente* (engl.: connected component). Seien zwei besetzte Oktanten Q und $V \in Okt^k$ gegeben, die nicht direkt benachbart sind, aber U als einen gemeinsamen, besetzten Nachbarn teilen. Es existiert somit eine *Sequenz* $Seq = [Q, U, V]$ zwischen Q, U und V , welche die Oktanten verbindet. Allgemein gilt folgende Definition: Zwei beliebige Oktanten Q und V gelten als verbunden, wenn mindestens die Sequenz $Seq = [Q = Okt_0, Okt_1, \dots, Okt_{n-1}, Okt_n = V]$ existiert, wobei ein Oktant Okt_i stets mit dem Oktanten $Okt_{i-1}, 1 \leq i \leq n$ benachbart ist.

Der Connected Components Algorithmus beginnt mit einem Oktanten U und seinen direkten Nachbarn Okt^k , die unmittelbar eine verbundene Komponente bilden. Jeder weitere besetzte Oktant $\notin Okt^k$, der aber direkter Nachbar eines Oktanten $\in Okt^k$ ist, gilt ebenfalls als mit der Komponente verbunden und wird zur Komponente hinzugefügt. Die auf diese Weise neu aufgenommenen Oktanten werden ebenfalls auf Verbundenheit mit ihren direkt benachbarten Oktanten überprüft. Sequentiell werden somit neue Oktanten zur Komponente hinzugefügt, bis kein Oktant mehr hinzugefügt werden kann, der nicht bereits in der Komponente enthalten ist. Die von den Oktanten der Komponente umschlossenen 3D-Punkte gelten damit ebenfalls als zusammenhängend.

Zwei separate Komponenten entstehen, wenn kein einziger besetzter Oktant mit minimaler Seitenlänge L_{min} als gemeinsamer Nachbar wenigstens eines ihrer Oktanten existiert. Die in den Oktanten der Komponenten positionierten 3D-Punkte gelten dann ebenfalls als nicht zusammenhängend. Die minimale Seitenlänge L_{min} definiert daher die maximale räumliche Distanz, die zwischen zwei Komponenten liegen darf, um sie noch als räumlich verbunden definieren zu können.

Zwei Parameter sind bei der Anwendung von CC für diese Arbeit von Bedeutung: 1. Die Tiefe D des Octrees und 2. die Anzahl von 3D-Punkten p_K , die mindestens in einer Komponente enthalten sein muss. In CloudCompare kann ein Octree eine maximale Tiefe von $D = 10$ besitzen. Die minimale Seitenlänge ist daher $L_{min} = L/2^{10}$. L_{min} wird daher von der Länge L des Wurzelraumes bestimmt und kann nicht beliebig klein sein. Für eine Länge der betrachteten Ausschnitte der Anbaureihen von $L = 5m$ und eine maximale Tiefe $D = 10$ beträgt L_{min} mindestens $5mm$. Insbesondere Trauben der TS-Erziehung liegen teilweise dichter als $5mm$ zusammen. Um diese dicht benachbarten Trauben zu trennen, muss die Länge L_{min} ungefähr $1mm$

betragen. Es wird daher sequentiell basierend auf A-priori-Wissen über den typischen Abstand von Trauben vorgegangen.

In einem ersten Schritt wird $D = 10$ gewählt, dass L_{min} ca. $5mm$ beträgt. Der Schwellwert $p_K > 0$ für die Anzahl an Punkten in einer Komponente gibt an, welcher der minimalen Punktzahl einer typischen Traube der Sorte und Erziehungsart entspricht. Das Connected Components Verfahren erzeugt daher Komponenten, die mindestens $5mm$ auseinanderliegen. Komponenten mit einer Anzahl an Punkten $\leq p_K$ werden dabei entfernt. Kleine Punktregionen von Fehlklassifikationen werden auf diese Weise gelöscht. Kleinere tatsächliche Trauben, nur teilweise rekonstruierte oder nur teilweise korrekt klassifizierte Trauben können jedoch ebenfalls gelöscht werden.

In einem zweiten Schritt werden diejenigen erzeugten Komponenten erneut betrachtet, deren Punktzahl einen empirisch ermittelten Maximalwert von p_{max} übersteigt. Eine Punktzahl $\geq p_{max}$ deutet auf die Fusion mehrerer Einzeltrauben zu einer einzigen Komponente hin. Diese Trauben liegen dichter als $5mm$ zusammen und konnten daher im ersten Schritt räumlich nicht separiert werden. CC wird daher nochmals auf diese Komponenten angewendet. Der Wurzelraum besitzt nun lediglich die Länge L der Komponente, die üblicherweise $50cm$ nicht überschreitet. L_{min} kann daher nun kleiner als $5mm$ gewählt werden. D wird nun so gewählt, dass L_{min} ca. $1mm$ beträgt, wobei p_K wie im ersten Schritt gewählt wird. Auf diese Weise können dicht benachbarte Trauben im Abstand $< 5mm$, die vorher zu einer einzigen Komponente gehörten, räumlich getrennt und als eigene Komponente entstehen.

Die mittels des CC-Verfahrens detektierten N-Komponenten $\in \mathcal{P}^T$ gelten zunächst als Traubenkandidaten $\mathcal{P}_i^{Tk} \in \mathcal{P}^T$ und konstituieren die Menge $\mathcal{P}^{Tk} = \{\mathcal{P}_1^{Tk} \dots \mathcal{P}_i^{Tk} \dots \mathcal{P}_N^{Tk}\}, i = 1 \dots N$. Unter Umständen verbleiben jedoch weiterhin einige Komponenten in $\mathcal{P}^{Tk}$, in denen mehrere, sich berührende Trauben vereint sind. Zusätzlich bleiben u. U. Komponenten bestehen, die keiner Trauben entsprechen aber akzeptiert wurden, weil sie $> p_K$ Punkte enthalten. Abbildung 4.26 illustriert eine solche Situation:



Abbildung 4.26: Durch das CC-Verfahren generierte Komponenten, umrahmt durch eine rote Bounding Box. Die Traubenkandidaten im unteren linken Teil des Bildes konnten nicht weiter räumlich separiert werden und werden zu einer einzigen Komponente zusammengefasst. Im unteren rechten Bild sind beispielhaft einige Fehlklassifikationen erkennbar, die nicht durch p_K gefiltert werden konnten. Im oberen Bereich befinden sich räumlich dicht zusammenliegende, aber individuelle Trauben, die korrekt in einzelne Komponenten unterteilt wurden.

Je nach Erziehungssystem und Reifestadium müssen unterschiedliche Werte für p_K und p_{max} für ein optimales Ergebnis gewählt werden (s. Kapitel 4.3.2 zur Detektion von Trauben und Beeren in beiden Reifestadien und Erziehungssystemen), da sich die Größe und damit die Punktzahl der Trauben darin jeweils unterscheiden. Die Werte sind dabei jeweilig empirisch zu ermitteln.

4.3.1.2 Detektion von Trauben und Beeren mit dem Berryfinder

Algorithmus 4.1 : findSpheres($\mathcal{P}_i^{Tk}$, Schwellwerte)

Data : Initialisiere Punktmenge $\mathcal{P}_i^{Tk} \in \mathcal{P}^{Tk}$; setze Schwellwerte**Result** : Menge der Beeren B , verbleibende Punkte in $\mathcal{P}_i^{Tk}$

```

1 while # Iterationen < Anzahl Punkte  $\in \mathcal{P}_i^{Tk}$  do
2   Wähle zufälligen Punkt aus  $\mathcal{P}_i^{Tk}$ ;
3   Schätze Kugel aus Punktnachbarschaft im Umkreis  $r_N$  um zufälligen Punkt;
4   Prüfe Kugelradius;
5   if Kugelradius innerhalb zu erwartender Größenordnung then
6     Bestimme Nachbarn im Kugelradius um Mittelpunkt der geschätzten Kugel;
7     Bestimme Anzahl und Anteil unterstützender Nachbarn der Kugelschätzung;
8     if Punktzahl und Anteil für Unterstützung groß genug then
9       Schätze Kugel erneut mit allen Nachbarn um Kugel;
10      Prüfe erneut den Anteil unterstützender Nachbarn für die zweite Kugelschätzung;
11      if Unterstützung groß genug then
12        Prüfe, ob die Kugel auf der Innenseite oder der Außenseite von  $\mathcal{P}_i^{Tk}$  liegt;
13        if Liegt auf Innenseite von  $\mathcal{P}_i^{Tk}$  then
14          Akzeptiere Kugel als legitime Beerenschätzung;
15          Füge Kugel zur Menge  $B$  hinzu;
16          Entferne Nachbarn um Kugelmittelpunkt aus  $\mathcal{P}_i^{Tk}$ ;
17          Iterationen = 0;
18          Beginne von Neuem;
19        else
20          Iterationen = Iterationen + 1;
21        end
22      else
23        Iterationen = Iterationen + 1;
24      end
25    else
26      Iterationen = Iterationen + 1;
27    end
28  else
29    Iterationen = Iterationen + 1;
30  end
31 end

```

Die Detektion von Beeren in den Traubenkandidaten ist komplex, da die eigentlich runde Geometrie der Beeren nicht immer völlig exakt rekonstruiert werden kann oder Beeren nicht völlig rund wachsen. Es wird daher ein eigener Algorithmus zur Zählung der Beeren kreiert, der vollen Zugriff für seine Adaption auf verschiedene Erscheinungsformen der Beeren ermöglicht. Er basiert auf A-priori-Wissen:

- a) Beeren sind annähernd kugelförmig. Die betrachteten Traubenkandidaten bestehen aus dicht benachbarten Kugelsegmenten, dem von der Kamera einsehbaren Teil der Traubenoberfläche.
- b) Die Kugelsegmente besitzen eine bekannte Größenspanne, entsprechend dem Reifegrad der Beeren bei der Bildaufnahme.
- c) Die räumlich nächstgelegene Kameraposition $\mathbf{p}_{cam}$ relativ den Punkten des Traubenkandidaten ist bekannt.

Der folgende Abschnitt beschreibt den Algorithmus *Berryfinder* zur Detektion und Größenbestimmung von Beeren in den Traubenkandidaten $\mathcal{P}^{Tk}$. Er basiert auf Kugelschätzungen und umfasst mehrere hierarchische

Evaluationsebenen, um festzustellen, ob eine Kugelschätzung einer Beeren entspricht. Zunächst wird die Funktion *findSpheres* beschrieben, die innerhalb des Berryfinders aufgerufen wird. Sie ist im Pseudocode 4.1 beschrieben.

Funktion *findSpheres*

Aus der Komponente $\mathcal{P}_i^{Tk}$ wird zunächst ein zufälliger Punkt $\mathbf{p}$ ausgewählt. Mittels eines kd-trees wird eine Punktnachbarschaft $\mathcal{P}^k$ im Radius r_N um $\mathbf{p}$ bestimmt. Da ein Traubenkandidat größtenteils aus Kugelsegmenten besteht, ist die Chance groß, dass die Punkte $\mathcal{P}^k$ einen Teil der Oberfläche eines Kugelsegments abbildet. Die Punkte in $\mathcal{P}^k$ werden anschließend zur Approximation einer Kugel $K1$ herangezogen. Die approximierten Kugel besitzt den Mittelpunkt $\mathbf{c}_{K1}$ und den Kugelradius r_{K1} . Danach wird die Kugelschätzung hierarchisch evaluiert:

1. **Zeile 2-5) – Radius r_K :** Die untersuchten Beeren besitzen eine Spanne an Radien $[r_{min}; r_{max}]$, die ihrem momentanen Reifegrad entspricht. Es wird geprüft, ob r_k innerhalb von $[r_{min}; r_{max}]$ liegt. Kugeln mit Radien außerhalb dieser Spanne werden verworfen.
2. **Zeile 6-11) – Unterstützung:** Bestimme die Nachbarn $\mathcal{P}^c$ um $\mathbf{c}_{K1}$ im Radius $r_{K1} + k_u$. Falls die Kugel einer Beere entspricht, werden alle Nachbarn um $\mathbf{c}_{K1}$ zur Beere gehören. Der Korrekturwert k_u wird aus zwei Gründen eingeführt:
 - a) die Nachbarnpunkte sind auch nach der Glättung mit CloudCompare mit kleinen zufälligen Abweichungen behaftet.
 - b) die Rekonstruktion der Beere weicht von der symmetrischen Kugelform ab (bspw. annähernd Ellipsen- oder Parabelförmig, abgeflachte Form).

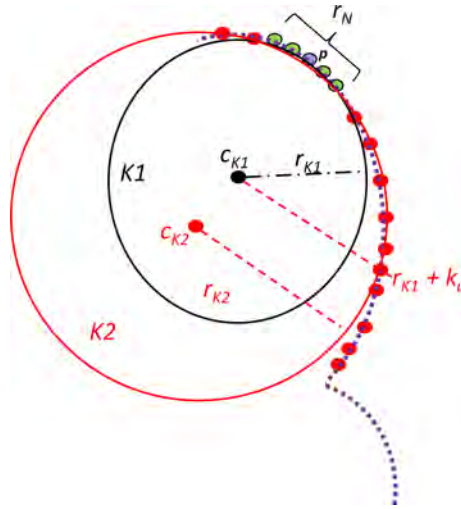


Abbildung 4.27: Bestimmung der Kugeln $K1$ (schwarzer Kreis) und $K2$ (roter Kreis) und Prüfung auf unterstützende Nachbarnpunkte. Gestrichelte violette Linie: 3D-Rekonstruktion einer Traube im Profil betrachtet. Lila Punkt: Zufällig ausgewählter Punkt $\mathbf{p}$. Grüne Punkte: Nachbarnpunkte $\mathcal{P}^k$. Rote Punkte: Nachbarnpunkte $\mathcal{P}^c$ um Kugelmittelpunkt $\mathbf{c}_{K1}$ im Radius $r_{K1} + k_u$.

Abbildung 4.27 stellt das geschilderte Vorgehen dar. Durch Einführen von k_u werden diese Phänomene berücksichtigt und möglichst alle Nachbarnpunkte bestimmt, die zur tatsächlichen Beere gehören. Es wird anschließend der Abstand jedes Nachbarnpunktes in $\mathcal{P}^c$ zu $\mathbf{c}_{K1}$ berechnet und deren Abstand zum Radius der Kugelschätzung verglichen. Punkte, deren Abstand zum Kugelmittelpunkt mehr als 10% von r_{K1} abweicht, werden als Ausreißer behandelt, alle anderen als unterstützende Punkte $\mathcal{P}^{in}$. Im Anschluss wird das Anzahlverhältnis

$$r_{in} = \frac{\#\mathcal{P}^{in}}{\#\mathcal{P}^c} \quad (4.11)$$

von $\mathcal{P}^{in}$ zu allen Nachbarn $\mathcal{P}^c$ berechnet. Die Kugel wird verworfen, wenn r_{in} und $\#\mathcal{P}^{in}$ gewählte Schwellwerte nicht erreichen. Der Schwellwert für die Anzahl unterstützender Nachbarpunkte $\#\mathcal{P}^{in}$ wird eingeführt, um Kugelschätzungen mit nur wenigen Nachbarn zu vermeiden. Als Schwellwert dienen empirische Werte der durchschnittlichen Punktzahl einer Beere.

Ist die Unterstützung groß genug, korrespondieren Kugelmittelpunkt und Kugelradius zunächst ausreichend mit den Daten. Die Kugel wird erneut mit allen Nachbarpunkten $\mathcal{P}^c$ approximiert, um die in a) und b) erläuterten Phänomene zu kompensieren. Die approximierte Kugel besitzt den Mittelpunkt $\mathbf{c}_{K2}$ und den Kugelradius r_{K2} . Bei der zweiten Kugelschätzung wird ein Radius in der Spanne von $[r_{min} - k_{min}; r_{max} + k_{max}]$ zugelassen, um Abweichungen von der Kugelgeometrie und natürliche Variationen in der erwarteten Größenentwicklung der Beeren zu berücksichtigen. Es wird erneut das Anzahlverhältnis r_{in} nach dem oben beschriebenen Verfahren anhand der Nachbarpunkte $\mathcal{P}^c$ und dem neuen Kugelmittelpunkt $\mathbf{c}_{K2}$ geprüft und für dieses erneut die im vorherigen Schritt definierten Schwellwerte für r_{in} und $\#\mathcal{P}^{in}$ zum Erreichen der nächsten Evaluationsebene genutzt.

3. **Zeile 12-18) – Position:** Die Außenseite der Traube wird als der Bereich definiert, der zwischen den Nachbarpunkten $\mathcal{P}^c$ und der zur Kugel nächstgelegenen Kameraposition $\mathbf{p}_{cam}$ liegt. Die Innenseite wird komplementär definiert als der Bereich vor den Nachbarpunkten und dem Außenbereich. Abbildung 4.28 illustriert die Vorgehensweise, sowie den Innen- und Außenbereich. Die Kugelmittelpunkte von Kugelschätzungen, die mit Beeren korrespondieren, befinden sich auf der Innenseite der Traube. Kugelschätzungen, die auf der Außenseite im Raum zwischen mehreren Beeren liegen, korrespondieren nicht mit Beeren und sollen abgelehnt werden.

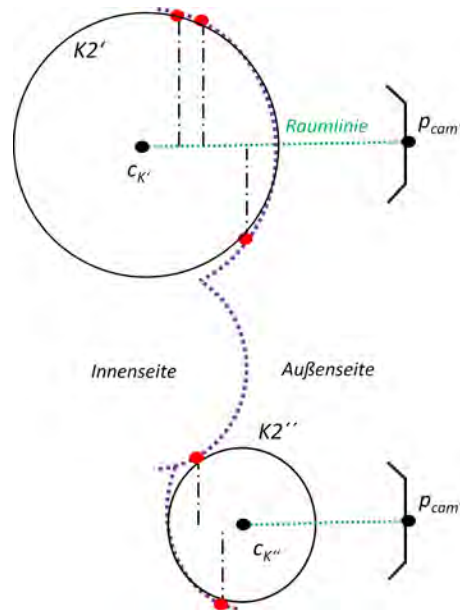


Abbildung 4.28: Definition der Innen- und Außenseite. Gestrichelte violette Linie: 3D-Rekonstruktion einer Traube im Profil betrachtet. Grüne gestrichelte Linie: Raumlinie zwischen Kugelmittelpunkt und Kameraposition. Rote Punkte: Jeweilige exemplarische Nachbarpunkte um die Kugelmittelpunkte der Kugeln $K2'$ und $K2''$ mit ihrer Orthogonalprojektion (punktiert-gestrichelte Linie) auf die Raumlinie zwischen den jeweilig nächsten Kamerapositionen $\mathbf{p}_{cam'}$ bzw. $\mathbf{p}_{cam''}$ und den jeweiligen Kugelmittelpunkten. $K2'$ liegt auf der Innenseite und wird akzeptiert, $K2''$ liegt auf der Außenseite und wird verworfen.

Zur Detektion von Kugeln auf der Außenseite wird eine Raumlinie über die nächstgelegene Kameraposition $\mathbf{p}_{cam}$ und den Kugelmittelpunkt $\mathbf{c}_K$ festgelegt. Nun wird geprüft, ob die Orthogonalprojektion aller Punkte in $\mathcal{P}^c$ auf die Raumlinie zwischen $\mathbf{p}_{cam}$ und $\mathbf{c}_K$ liegt. Liegen mehr als 70% aller Projektionen auf der Raumlinie zwischen diesen Punkten, befindet sich der Kugelmittelpunkt $\mathbf{c}_K$ auf der Innenseite der Traube und wird akzeptiert, andernfalls abgelehnt.

Die gefundene Kugel gilt zunächst als detektierte Beere und wird in der Menge B gespeichert. Die Nachbarpunkte $\mathcal{P}^c$ werden anschließend aus $\mathcal{P}_i^{Tk}$ entfernt.

Die Funktion `findSpheres` startet anschließend von Neuem, bis $\mathcal{P}_i^{Tk}$ entweder keine Punkte mehr enthält oder die Anzahl der Iterationen der Anzahl der verbliebenen Punkte in $\mathcal{P}_i^{Tk}$ entsprechen, ohne dass eine weitere Kugelschätzung als Beere akzeptiert wurde. Auf diese Weise werden möglichst viele Punkte als Quellpunkt für eine Kugelschätzung fungieren.

Hauptprogramm Berryfinder

Das Hauptprogramm Berryfinder wird im Pseudocode 4.2 und ab Punkt 4 geschildert. Im Berryfinder wird die Funktion `findSpheres` zweimal aufgerufen, Traubenkandidaten über einen Schwellwert für die minimal detektierte Anzahl von Beeren auf Korrespondenz mit tatsächlichen Trauben evaluiert und die Beeren auf Überlappung geprüft:

Algorithmus 4.2 : BerryFinder

Data : Initialisiere Traubenkandidat $\mathcal{P}_i^{Tk} \in \mathcal{P}^{Tk}$

Result : $\mathcal{P}_i^{Tk}$ als Traube akzeptiert; Anzahl und Größe der Beeren in $\mathcal{P}_i^{Tk}$;

```

1 foreach  $\mathcal{P}_i^{Tk} \in \mathcal{P}^{Tk}$  do
2   findSpheres( $\mathcal{P}_i^{Tk}$ , Schwellwerte hoch) Result : Menge  $B_1$ , Verbleibende Punkte in  $\mathcal{P}_i^{Tk}$  ;
3   if Anzahl der Beeren in  $B_1 > min_B$  then
4     findSpheres(Verbleibende Punkte in  $\mathcal{P}_i^{Tk}$ , Schwellwerte niedrig) Result : Menge  $B_2$ 
5     Menge  $B = B_1 \cup B_2$ ; Akzeptiere  $\mathcal{P}_i^{Tk}$  als Traube;
6     foreach Kugel  $\in B$  do
7       Wähle sequentiell nächste Kugel aus  $B$ ;
8       Prüfe, ob sich die Kugel zu einem gewissen Grad mit anderen Kugeln überlappt;
9       if Überlappung zu groß then
10        Wähle aus allen sich überlappenden Kugeln diejenige
11        mit der größten Unterstützung seiner Nachbarn im Kugelradius
12        um den Kugelmittelpunkt aus;
13      else
14        Beginne von Neuem
15      end
16    end
17  else
18    Nicht genügend Beeren detektiert. Verwerfe  $\mathcal{P}_i^{Tk}$  als Traube
19    Beginne von Neuem mit  $\mathcal{P}_{i+1}^{Tk}$ 
20  end
21 end
```

4. **Zeile 1-3, 17-19) – 1. Aufruf von `findSpheres`:** Der erste Aufruf übergibt der Funktion `findSpheres` einen Traubenkandidaten $\mathcal{P}_i^{Tk} \in \mathcal{P}^{Tk}$ und detektiert Beeren B_1 zunächst mit hohen Schwellwerten für die einzelnen Evaluationsebenen. Dies sichert die Detektion von möglichst exakt rekonstruierten, kugelförmigen Geometrien und erhöht die Wahrscheinlichkeit, dass die Kugelschätzung einer Beere entspricht. Werden mehr als min_B Beeren gefunden, wird der Traubenkandidat als Traube akzeptiert, andernfalls als Traubenkandidat verworfen.

Wird der Traubenkandidat verworfen, findet für ihn keine weitere Suche nach Beeren statt. Dieses Vorgehen dient dazu Traubenkandidaten in $\mathcal{P}_i^{Tk}$ zu identifizieren, die fälschlicherweise als **Traube** klassifizierte Punkte der Klasse **Rest** enthalten. Diese fehlklassifizierten Punktregionen enthalten üblicherweise nicht die typischen Kugelgeometrien, aus denen Trauben bestehen. Es ist daher für eine in diesen Punktregionen geschätzte Kugel schwieriger, die hoch gesetzten Schwellwerte von `findSpheres` für die einzelnen Evaluationsebenen zu erreichen und als legitime Beere akzeptiert zu werden. Dadurch

wird der Schwellwert $\min_B$ schwieriger erreicht und Traubenkandidaten mit falsch klassifizierten Punkten gefiltert. Nur teilweise rekonstruierte, tatsächliche Trauben in $\mathcal{P}^{T_k}$ können auf diese Weise u. U. fälschlicherweise entfernt werden, wenn zu wenige tatsächliche Beeren im Traubenkandidaten enthalten sind.

5. **Zeile 4-5) – 2. Aufruf von *findSpheres*:** Im 1. Aufruf von *findSpheres* wurden hohe Schwellwerte für die Evaluationsebenen definiert. Dadurch wurden nur dort Kugeln geschätzt, wo die Kugelgeometrie gut rekonstruiert worden ist. Die zur Schätzung verwendeten Punkte wurden dabei aus der Punktmenge im betrachteten Traubenkandidaten $\mathcal{P}_i^{T_k}$ entfernt. Der zweite Aufruf sucht in den verbliebenen Punkten in $\mathcal{P}_i^{T_k}$ erneut nach Kugeln und speichert diese in der Menge B_2 . Dabei werden niedrigere Schwellwerte als im 1. Aufruf angesetzt. Zu niedrige Schwellwerte im 1. Aufruf von *findSpheres* hätten zu einer erhöhten Anzahl von Kugeln geführt, die nicht mit tatsächlichen Beeren korrespondieren, was die Identifikation fehlklassifizierter Punktregionen erschweren würde.

Der 2. Aufruf mit niedrigeren Schwellwerten dient dazu, Beeren zu detektieren, deren rekonstruierte Geometrie nicht exakt kugelförmig ist, die aber dennoch höchstwahrscheinlich tatsächlichen Beeren entsprechen.

Zusammenfassend dient der 1. Aufruf zum Ausschluss falscher Traubenkandidaten, der 2. Aufruf zur möglichst vollständigen Detektion aller vorhandenen Beeren im akzeptierten Traubenkandidaten. Nachdem keine Beeren mehr detektiert wurden, werden B_1 und B_2 zur Menge B vereint.

6. **Zeile 6-15) – Überlappung:** Die Beeren in B werden zuletzt auf gegenseitige Überlappung geprüft. Beeren wachsen dicht benachbart und üben Druck aufeinander aus, welcher die Beeren leicht deformieren kann. Der nicht sichtbare Teil einer Beere weicht daher höchstwahrscheinlich von der Kugelform ab. Dieser Umstand wird durch Gestattung eines gewissen Überlappungsgrades von 30% des eigenen Beerenvolumens berücksichtigt.

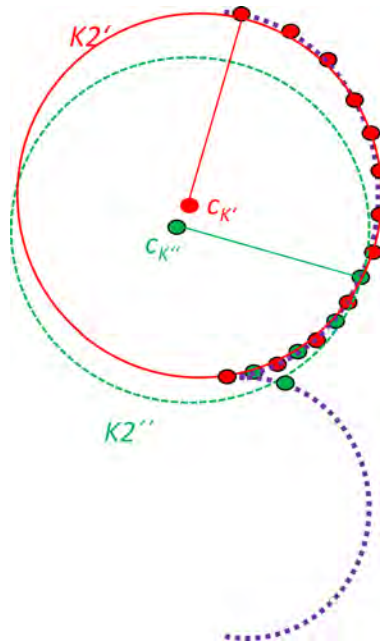


Abbildung 4.29: Zwei Beeren K' (rot) und K'' (grün) aus der Menge B mit unterstützenden Nachbarnpunkten in rot bzw. grün. Gestrichelte violette Linie: 3D-Rekonstruktion einer Traube im Profil betrachtet. Rote Punkte: Nachbarnpunkte Kugel von K' . Grüne Punkte: Nachbarnpunkte von K'' . Kugel K' wird als Lösungsvorschlag akzeptiert, da sie besser mit den Punkten korrespondiert. Kugel K'' hingegen besitzt 1. Nachbarnpunkte auf einer benachbarten Beere und 2. weniger unterstützende Punkte.

Unter Umständen werden mehrere Kugeln als Kandidat für die gleiche tatsächliche Beere geschätzt. Abbildung 4.29 illustriert die Situation. Diese resultiert aus Abweichungen von der Kugelgeometrie (s.

Punkt 2b). Nahe Punkte um c_{K2} , die tatsächlich zur selben Beere gehören, aber für die Kugelschätzung nicht in $\mathcal{P}^c$ enthalten waren, verbleiben für die nächste Iteration in $\mathcal{P}_i^{Tk}$ und ermöglichen damit eine weitere Kugelschätzung nahe einer Kugelschätzung aus einer vorausgehenden Iteration von *findSpheres*. Kugeln, die sich stärker als 30% überlappen, sollen daher auf ihre Korrespondenz mit den Daten geprüft und die Kugel mit der höchsten Übereinstimmung als Beere akzeptiert werden.

Seien beispielhaft zwei sich überlappende Kugeln $K2'$ und $K2''$ gegeben. Ihre Nachbarn im eigenen Kugelradius $r_{K2'}$ bzw. $r_{K2''}$ um $c_{K2'}$ bzw. $c_{K2''}$ werden bestimmt und ihre Unterstützung für die Kugelschätzung nach Formel 4.11 in Punkt 2 des Pseudocodes 4.1 zu *findSpheres* geprüft. Die Kugelschätzung mit einem Anzahlverhältnis r_{in} größer als ein gewählter Schwellwert wird als bessere Kugelschätzung akzeptiert. Übertreffen beide Kugelschätzungen diesen Wert, wird die Kugel mit der größeren Anzahl unterstützender Punkte $\#\mathcal{P}^{in}$ gewählt. Erreicht keine Kugel den gesetzten Schwellwert, wird diejenige mit der größeren r_{in} gewählt.

Falls sich die akzeptierte Kugel mit weiteren Kugeln über den erlaubten Schwellwert hinaus überlappt, werden diese erneut paarweise nach dem beschriebenen Verfahren geprüft und die beste Lösung ausgewählt. Alle Fälle von Überlappungen in B werden geprüft. Auf diese Weise kann aus mehreren Lösungsvorschlägen die beste Kugelschätzung ausgesucht werden. In B verbleiben dadurch jeweils nur die Kugeln, welche am besten zu den Daten passen. Die Anzahl der verbleibenden Kugeln in B gelten als geschätzte Anzahl der sichtbaren Beeren für den Traubenkandidaten $\mathcal{P}_i^{Tk}$. Die Suche nach Beeren ist damit für diesen Kandidaten abgeschlossen.

In jedem Traubenkandidaten $\mathcal{P}_i^{Tk} \in \mathcal{P}^{Tk}$ wird mittels des Berryfinders nach Beeren gesucht und Traubenkandidaten auf die vorherig beschriebene Weise entweder als Traube akzeptiert oder verworfen. Dies geschieht, bis alle Traubenkandidaten geprüft wurden.

Ergebnis nach der Prüfung aller Traubenkandidaten ist die Anzahl der detektierten, sichtbaren Trauben und Beeren, sowie deren geschätzte Beerengröße:

- Die Anzahl der nach Durchlauf des Berryfinders akzeptierten Traubenkandidaten mit mehr als min_B Beeren gelten als Anzahl der für die Kameras sichtbaren Trauben.
- Die Summe aller Beeren in den akzeptierten Traubenkandidaten gilt als Anzahl der für die Kameras sichtbaren Beeren.
- Der Durchmesser der geschätzten Kugeln gilt als Schätzung der jeweiligen Beerengröße.

4.3.1.3 Evaluation des Berryfinder

Bei der Untersuchung des Berryfinders sollen folgende Fragen beantwortet werden:

1. Wie viele Kugelschätzungen entsprechen tatsächlichen Beeren?
2. Wie viele vorhandene Beeren werden detektiert?
3. Wie präzise ist die Bestimmung der Beerenanzahl bei wiederholter Suche?
4. Wurde die Beerengröße korrekt bestimmt?

Evaluation anhand einer hochaufgelöst rekonstruierten Traube

Zur Beantwortung dieser Fragen werden zunächst hochaufgelöste Referenzdaten einer künstlichen Weintraube mit dem Linienscanner Perceptron v5 (Perceptron Scan Works V5, Perceptron Inc., Plymouth, MI, USA) erstellt, der mit dem Messarm Romer Infinite 2.0 (Hexagon Metrology Services Ltd., London, UK) gekoppelt wird [Dupuis, 2017]. Mit dem Linienscanner wird die künstliche Weintraube mit einem durchschnittlichen Punkt-zu-Punkt-Distanz von $14\mu m$ und einer Punktgenauigkeit von $45\mu m$ rekonstruiert.

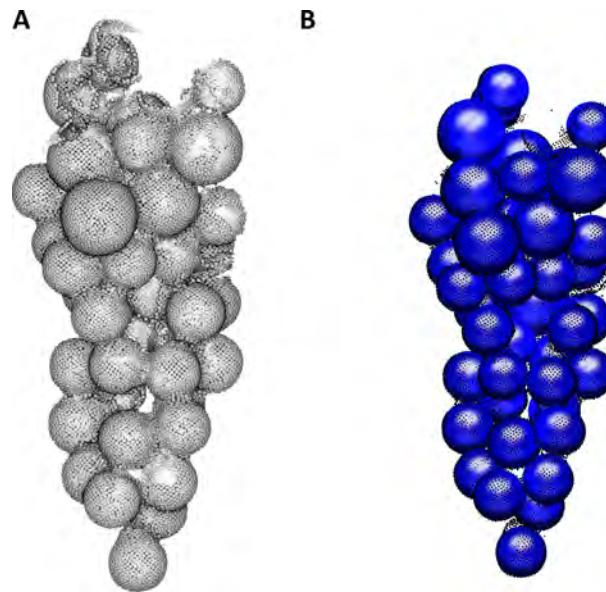


Abbildung 4.30: **A:** Die mit dem Perceptron v5 erstellte Punktwolke der künstlichen Weintraube. **B:** Die durch den Berryfinder detektierten Beeren. Partiiell rekonstruierte Beeren genügen zur Approximation einer Kugel, aus Rose et al. [2016].

Die tatsächliche Anzahl der Beeren wurde manuell gezählt. Der Umfang von 20 Beeren wurde mit einem Messschieber gemessen und betrug $15\text{--}20\text{mm}$. Mittels des Berryfinders werden Beeren in der Punktwolke der Weintraube gesucht. Für den Berryfinder wird im 1. Aufruf der Funktion *findSpheres* ein Nachbarschaftsradius $r_N = 3\text{mm}$, sowie die Schwellwerte $r_{in} = 80\%$ und $\#P^{in} = 60$ angesetzt. Für den 2. Aufruf von *findSpheres* werden $r_N = 4\text{mm}$, sowie die Schwellwerte $r_{in} = 60\%$ und $\#P^{in} = 40$ gewählt. Diese Werte wurden durch Ausprobieren und visuelle Überprüfung als optimal bestimmt.

Visuell wird untersucht, wie viele Kugeln mit tatsächlichen Beeren korrespondieren. Aus den 20 vermessenen Beeren wird außerdem anhand des manuell gemessenen Radius' die Standardabweichung des geschätzten Radius berechnet. Abbildung 4.30 zeigt die Punktwolke der künstlichen Weintraube und die mit dem Berryfinder detektierten Beeren.

Alle Beeren wurden vom Berryfinder detektiert. Alle Kugelschätzungen korrespondieren mit tatsächlichen Beeren. Im Vergleich des geschätzten Radius der 20 vermessenen Beeren mit den manuell gemessenen Radien ergibt sich eine Standardabweichung von $0,5\text{mm}$.

Evaluation anhand einer aus Bildern des PHENObots rekonstruierten Traube

Der Berryfinder wird im Weiteren anhand einer aus den PHENObot Bildern rekonstruierten Weintraube aus der TS-Anbaureihe im BBCH89 Reifestadium untersucht. Es soll dadurch untersucht werden, welche Genauigkeit der Berryfinder bei Daten erreicht, die unter erschwerten Feldbedingungen mittels SfM/MVS erstellt wurden. Die Weintraube umfasst insgesamt 43 Beeren, die manuell in der Punktwolke gezählt wurden. Für den Berryfinder wird im 1. Aufruf der Funktion *findSpheres* ein Nachbarschaftsradius $r_N = 2\text{mm}$, sowie die Schwellwerte $r_{in} = 75\%$ und $\#P^{in} = 50$ angesetzt. Für den 2. Aufruf von *findSpheres* werden $r_N = 4\text{mm}$, sowie die Schwellwerte $r_{in} = 60\%$ und $\#P^{in} = 30$ gewählt. Diese Werte wurden durch Ausprobieren und visuelle Überprüfung als optimal bestimmt.

Abbildung 4.31 (A) zeigt die Punktwolke der Weintraube. In (B) ist das Zwischenergebnis des 1. Aufruf von *findSpheres* im Berryfinder mit hohen Schwellwerten dargestellt. Grüne Punkte stellen unterstützende Nachbarpunkte dar. Rote Punkte stellen die zur Kugelschätzung genutzten Punkte dar. Kugeln, die grüne Punkte lediglich auf der unteren Seite besitzen (beispielhaft mit roten Pfeile gekennzeichnet), liegen auf der Außenseite und werden im weiteren Verlauf entfernt. Abbildung (C) zeigt zufällig gewählte Quellpunkte zur

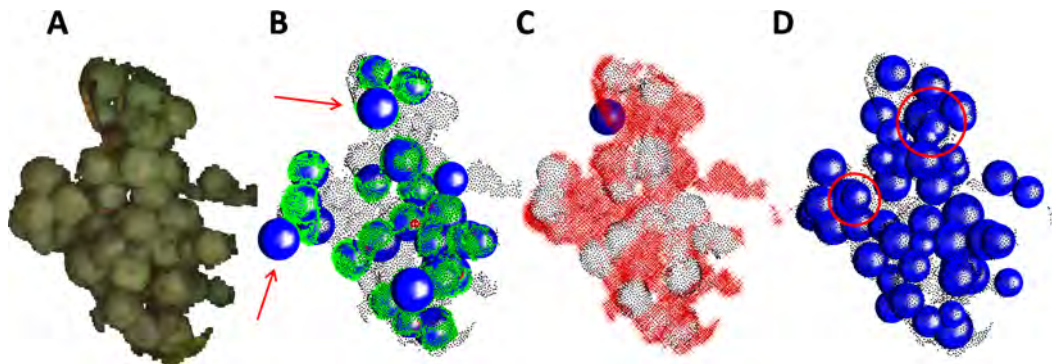


Abbildung 4.31: **A:** Die Punktwolke der Weintraube. **B:** Zwischenergebnis des 1. Aufruf von *findSpheres* im Berryfinder mit hohen Schwellwerten. **C:** Zufällig gewählte Quellpunkte zur Kugelschätzung im 1. Aufruf von *findSpheres* mit hohen Schwellwerten. **D:** Zwischenergebnis nach dem 2. Aufruf von *findSpheres* im Berryfinder.

Kugelschätzung im 1. Aufruf von *findSpheres* mit hohen Schwellwerten (rote Kreuze). Die Dichte gewählter Quellpunkte ist dort am höchsten, wo die Geometrie von einer exakten Kugelform abweicht und Kugelschätzungen schwerer akzeptiert werden können. Die geschätzten Kugeln aus (B) korrespondieren mit den Orten, wo aufgrund der gut rekonstruierten Kugelgeometrie nur wenige Quellpunkte zur Akzeptanz einer Kugel als Beere nötig waren. In (D) wird das Zwischenergebnis nach dem 2. Aufruf von *findSpheres* im Berryfinder gezeigt. Nach dem 2. Aufruf wurden Kugeln als Beeren akzeptiert, deren Nachbarpunkte keine exakte Kugelform abbilden, aber dennoch mit tatsächlichen Beeren korrespondieren. Einige Kugelschätzungen überlappen sich über das erlaubte Maß hinaus (rot umkreist). Diese werden im Anschluss bei der Prüfung auf Überlappung entfernt, wobei die Kugel mit der größten Unterstützung als legitime Beere akzeptiert wird.

Der Berryfinder wählt die Quellpunkte zur Kugelschätzung zufällig im Traubenkandidaten aus. Die Präzision der Zählung der Beerenanzahl soll daher geprüft werden. Um die Standardabweichung der geschätzten Beerenanzahl zu berechnen wird der Berryfinder 20 mal auf dieselbe betrachtete Weintraube angewandt.

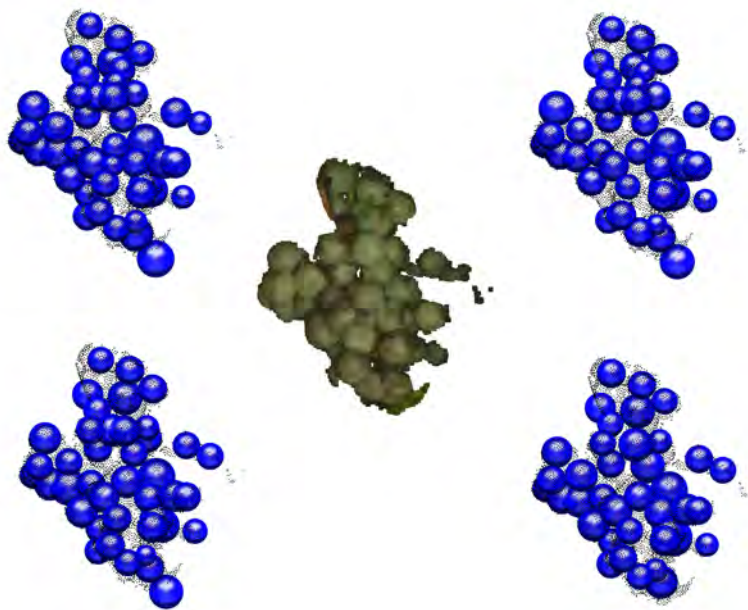


Abbildung 4.32: Die Punktwolke der Traube befindet sich in der Mitte. Vier beispielhafte Endergebnisse der 20 Anwendungen des Berryfinder sind in den Ecken dargestellt, aus Rose et al. [2016].

Abbildung 4.32 illustriert vier beispielhafte Endergebnisse der 20 Anwendungen des Berryfinders auf die Punktwolke der Weintraube. Die Standardabweichung bei 20 Anwendungen betrug 1,5 Beeren, bei einem Mittelwert von 42 detektierten Beeren bei einer Anzahl von 43 tatsächlich vorhandenen Beeren. Über visuelle Kontrolle wurde verifiziert, dass der Großteil der Beeren bei jeder Anwendung erneut detektiert wurde. Maximal zwei der detektierten Kugeln entsprachen dabei keinen tatsächlichen Beeren. Diese werden nur an restlichen 3D-Punkten von Blättern oder Ästen detektiert. Alle übrigen Kugeln entsprachen tatsächlichen Beeren.

Einige Beeren wurden jedoch nicht immer detektiert. Diese liegen vornehmlich größtenteils von anderen Beeren verdeckt im tieferen Bereich der Traube, weshalb nur wenige 3D-Punkte die Beerenoberfläche abbilden. Aufgrund der zufälligen Suche nach einem ersten Quellpunkt für die Kugelschätzung können diese 3D-Punkte unter Umständen in einer vorherigen Iteration als unterstützende Nachbarpunkte einer benachbarten Beere gelten. Sie werden dann aus der Punktmenge der Traube entfernt, weshalb die verbleibenden 3D-Punkte der Beere dann u. U. nicht mehr zur Kugelapproximation ausreichen.

Da Sorte, Reifestadium und Erziehungssystem maßgeblich Größe der Beeren und die Genauigkeit der Rekonstruktion der Beeren beeinflussen, sind die Parameter des Berryfinders entsprechend für jede Erscheinungsform der Trauben anzupassen.

4.3.1.4 Zusammenfassung

In den letzten drei Unterkapiteln wurden das CC-Verfahren sowie der Berryfinder als kombiniertes Verfahren zur Bestimmung der Anzahl sichtbarer Trauben, Beeren und der Beerengröße aus den Punkten der Klasse **Traube** vorgestellt und evaluiert.

Das CC-Verfahren teilt räumlich zusammenhängende Punktregionen der Klasse **Traube** in Komponenten ein, die als Traubenkandidaten an den Berryfinder weitergegeben werden. Dabei werden zusätzlich Punktregionen gelöscht, die eine nutzerdefinierte Punktzahl p_K nicht erreichen. Dabei handelt es sich größtenteils um fehlklassifizierte Punktregionen, jedoch können dadurch ebenfalls 3D-Punkte gelöscht werden, die tatsächliche Trauben abbilden. Des Weiteren können mehrere Trauben in einer Komponente zusammengefasst sein, die sehr dicht zusammen liegen und sich räumlich nicht trennen lassen oder es können Komponenten bestehen bleiben, die falsch klassifizierte Punkte enthalten.

Im Berryfinder werden Kugelapproximationen zur Detektion von Beeren genutzt. Bei der Evaluierung des Berryfinders zeigte sich, dass der Erfolg der Beerendetektion davon abhängt, wie exakt die Geometrie der Beeren rekonstruiert wurde und wie dicht die Beeren zusammen liegen. Die Schätzung der Größe der Beeren erreicht für die hochaufgelöste Punktwolke der künstlichen Weintraube eine Standardabweichung von 0,5mm bzgl. der Referenzdaten. Alle detektierten Beeren korrespondierten mit tatsächlichen Beeren und alle Beeren konnten detektiert werden. Die Standardabweichung der bestimmten Beerenzahl bei Daten, die unter Feldbedingungen erhoben wurden, beträgt 1,5 Beeren, bei einer mittleren geschätzten Beerenzahl von 42 und einer tatsächlichen Beerenzahl von 43. Die Standardabweichung erklärt sich aus der nicht immer völlig exakt rekonstruierten Geometrie der Beeren bei Erhebung unter Feldbedingungen und natürlichen Abweichungen der Beeren von der Kugelform. Die Ergebnisse weisen den Berryfinder als geeignetes Werkzeug zur Detektion von Beeren und zur Schätzung ihrer Größe aus.

Der Berryfinder wird genutzt, um in den im CC-Verfahren identifizierten Traubenkandidaten Beeren zu finden. Traubenkandidaten, in denen im 1. Aufruf von *findSpheres* weniger als min_B Beeren gefunden wurden, werden dabei als mögliche Trauben verworfen. Dies dient der Steigerung der Präzision der Traubendetektion. Unter Umständen können auf diese Weise tatsächliche Trauben entfernt werden, die bspw. aufgrund hoher Verdeckung oder durch die Klassifikation nur unvollständig vorhanden sind und zu wenige Beeren im Traubenkandidaten für eine Akzeptanz als Traube verbleiben.

4.3.2 Evaluation der Auswertemethode

Ziel dieser Arbeit ist es, die Anzahl **aller** Beeren und Trauben, die Beerengröße und daraus den Ertrag zu schätzen. Generelles Problem bei Verwendung einer optischen Sensorik ist jedoch, dass 1. höchstwahrscheinlich nicht alle sichtbaren Beeren und Trauben detektiert werden können, es 2. verdeckte, stets unsichtbare

Beeren und Trauben geben wird und 3. nicht alle detektierten Kugeln bzw. Komponenten tatsächlichen Beeren bzw. Trauben entsprechen. Ohne eine Schätzung der Gesamtanzahlen von Beeren und Trauben kann keine genaue Ertragsschätzung stattfinden.

In diesem Kapitel wird die Auswertemethode evaluiert. Folgende Fragen sollen beantwortet werden:

1. Wie viele Beeren und Trauben sind für die Kameras sichtbar?
2. Wie vollständig werden die sichtbaren Beeren und Trauben detektiert?
3. Wie viele der sichtbaren Beeren und Trauben werden **nicht** detektiert und was ist die Ursache dafür?
4. Wie viele detektierte Beeren und Trauben entsprechen tatsächlichen Beeren und Trauben?
5. Entspricht die geschätzte Beerengröße der tatsächlichen Beerengröße?

4.3.2.1 Zusammensetzung der Gesamtanzahlen der Beeren und Trauben

Zum weiteren Verständnis muss die Zusammensetzung der Gesamtanzahlen der Beeren und Trauben erläutert werden, die sich aufgrund der Verwendung einer optischen Sensorik ergibt. Die Gesamtanzahlen von Beeren bzw. Trauben in einer Anbaureihe seien mit $\mathcal{N}^B$ bzw. $\mathcal{N}^T$ bezeichnet. Der für die Kameras sichtbare Teil der Beeren bzw. Trauben mit $\mathcal{V}^B$ bzw. $\mathcal{V}^T$ bezeichnet. Es ist zu erwarten, dass die Auswertemethode nicht alle $\mathcal{V}^B$ bzw. $\mathcal{V}^T$ detektiert und dass nicht alle detektierten Beeren/Trauben auch tatsächlich Beeren/Trauben darstellen. Des Weiteren werden einige Beeren/Trauben so tief im Blattwerk liegen, dass sie für die Kameras stets unsichtbar sind.

Es sei:

- $\mathcal{D}^{B/T}$ (detected): Anzahl der von der Auswertemethode detektierten Beeren bzw. Trauben.
- $\mathcal{X}^{B/T}$ (excess): Anzahl der als Beere bzw. Traube detektierten Kugeln bzw. Komponenten, die keine tatsächlichen Beeren bzw. Trauben darstellen und um die $\mathcal{D}^{B/T}$ zu hoch ist.
- $\mathcal{M}^{B/T}$ (missing): Anzahl der für die Kameras sichtbaren, aber **nicht** mittels der Auswertemethode detektierten Beeren bzw. Trauben, um die $\mathcal{D}^{B/T}$ zu gering ist.
- $\mathcal{O}^{B/T}$ (occluded): Anzahl der für die Kameras **stets unsichtbaren** Beeren bzw. Trauben, die während der Datenerhebung immer verdeckt waren und daher nicht detektiert werden können¹.
- $\mathcal{N}^{B/T}$ (true number): Gesamtanzahl der Beeren bzw. Trauben in der Anbaureihe.
- $\mathcal{V}^{B/T}$ (true visible): Anzahl der für die Kameras sichtbaren Beeren bzw. Trauben

wobei $^{B/T}$ Beeren bzw. Trauben bezeichnet. Es ergeben sich folgende Beziehungen:

$$\mathcal{D}^{B/T} = \mathcal{V}^{B/T} + \mathcal{X}^{B/T} - \mathcal{M}^{B/T} \quad (4.12)$$

$$\mathcal{V}^{B/T} = \mathcal{D}^{B/T} - \mathcal{X}^{B/T} + \mathcal{M}^{B/T} \quad (4.13)$$

$$\mathcal{O}^{B/T} = \mathcal{N}^{B/T} - \mathcal{V}^{B/T} \quad (4.14)$$

¹Da in diesem Kapitel die inhärenten Eigenschaften der Auswertemethode behandelt werden, entfällt an dieser Stelle noch eine Betrachtung von $\mathcal{O}^T$. Eine Untersuchung von $\mathcal{O}^T$ erfolgt stattdessen in Kapitel 5, in dem die Schätzung des Ertrages aus den Ertragsparametern behandelt wird.

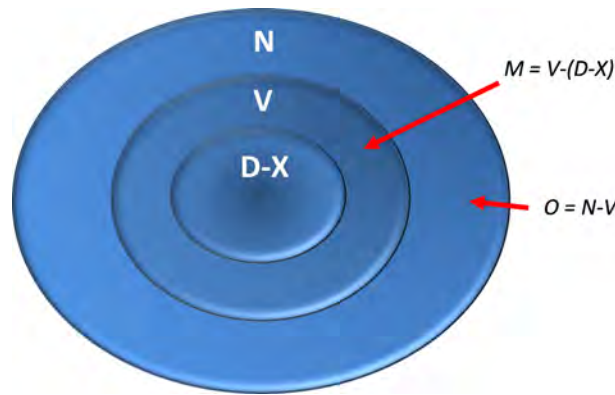


Abbildung 4.33: Die Gesamtanzahl $\mathcal{N}$ der Beeren bzw. Trauben setzt sich aus $\mathcal{D}$, $\mathcal{X}$, $\mathcal{M}$ und $\mathcal{O}$ zusammen.

Abbildung 4.33 illustriert die Beziehung der Anzahlen zueinander. Die jeweilige Gesamtanzahl ergibt sich aus:

$$\mathcal{N}^{B/T} = \mathcal{D}^{B/T} - \mathcal{X}^{B/T} + \mathcal{M}^{B/T} + \mathcal{O}^{B/T} \quad (4.15)$$

4.3.2.2 Vorgehensweise zur Evaluation der Auswertemethode

Um die obigen Fragen zu beantworten, werden in den $5m$ langen, bereits klassifizierten Ausschnitten der Anbaureihen (s. Kapitel 4.2.3) die sichtbaren Beeren und Trauben detektiert, sowie die Größe der Beeren geschätzt. Die $5m$ langen Ausschnitte werden dazu in jeweils $1m$ Abschnitte unterteilt und einzeln untersucht. Die Ergebnisse werden dabei je Erziehungsmethode und Reifestadium betrachtet. Dadurch wird ein Eindruck über die Funktionsweise der Auswertemethode und sich eventuell ergebende Unterschiede bei der Detektion der Beeren und Trauben beider Reifestadien und Erziehungssysteme gewonnen.

Bewertung der Detektion mit Recall und Precision

Der Erfolg bei der Detektion der sichtbaren Beeren und Trauben wird über die bereits bei der Klassifikation der Punktwolken genutzten Metriken Recall und der Precision bezogen auf $\mathcal{D}^{B/T}$ evaluiert (Formeln 4.8 und 4.9 in Kapitel 4.2). Recall und Precision sind Metriken, welche die Vollständigkeit bzw. Präzision einer Klassifikation quantifizieren. Bezogen auf die Detektion wird der Recall von $\mathcal{M}^{B/T}$ bestimmt (False Negatives), die Precision von der Anzahl $\mathcal{X}^{B/T}$, die nicht mit Beeren bzw. Trauben korrespondieren (False Positives). Da die Detektion im Grunde eine Klassifikation darstellt, die unterschiedlich vollständig, falsch oder richtig sein kann (bspw. kann eine Kugel zur Beerenzählung beitragen, die keiner tatsächlichen Beere entspricht) sind Recall und Precision angemessene Werkzeuge, um Vollständigkeit und Präzision der Beeren- und Traubendetektion zu bewerten. Sind $\mathcal{X}^{B/T}$ und $\mathcal{M}^{B/T}$ beispielsweise gleich Null, dann gilt nach Formel 4.13 $\mathcal{V}^{B/T} = \mathcal{D}^{B/T}$. Recall und Precision betragen in diesem Fall 100%.

Um detaillierte Kenntnisse über die Funktionsweise der Auswertemethode zu erlangen, wird des Weiteren untersucht, warum Trauben von der Auswertemethode nicht detektiert werden.

Referenzwerte für die Traubenanzahl

Um Recall und Precision zu berechnen werden die True Positives, False Negatives und False Positives der Detektion bestimmt. Die geschieht durch manuelle Zählung der Trauben in den Bildern und Punktwolken in jedem Schritt der Auswertemethode und durch Vergleich von $\mathcal{V}^T$, $\mathcal{D}^T$, $\mathcal{X}^T$ und $\mathcal{M}^T$. Eine Traube gilt als sichtbar und wird zu $\mathcal{V}^T$ gezählt, falls sie **mindestens zum Teil in mindestens einem Bild** abgebildet ist.

Die True Positives werden wie folgt gezählt:

- Eine Traube, die korrekt in einer einzigen Komponente enthalten ist, zählt als ein True Positive. Dabei ist es ausreichend, dass lediglich ein Teil der Traube in der Komponente enthalten ist, um die Traube als gezählt zu betrachten.

Auf den Bildern sichtbare, aber nicht detektierte Trauben ($\mathcal{M}^T$) zählen als False Negatives. Trauben werden aus einem der folgenden Gründe nicht detektiert:

1. Nicht rekonstruiert: Die Traube ist auf den Bildern sichtbar, wurde jedoch von Pix4DMapper nicht rekonstruiert. Sie ist nicht in der Punktwolke $\mathcal{P}$ enthalten.
2. Falsch klassifiziert: Die Traube wurde rekonstruiert, jedoch wurden alle Punkte der Traube falsch klassifiziert. Sie ist nicht in der Punktwolke $\mathcal{P}^T$ der als **Traube** klassifizierten Punkte enthalten.
3. Punktzahl $< p_K$: Die Komponente enthält eine Traube, erreichte während des CC-Verfahrens aber nicht die nötige Punktzahl p_K und wurde entfernt. Sie ist nicht in den Traubenkandidaten $\mathcal{P}^{Tk}$ enthalten.
4. Beerenanzahl $< \min_B$: Im Traubenkandidaten $\mathcal{P}_i^{Tk}$ wurden weniger als $\min_B$ Beeren gefunden. Der Traubenkandidat wurde daher verworfen, obwohl er eine tatsächliche Traube darstellte.
5. Nicht trennbar: Der Traubenkandidat $\mathcal{P}_i^{Tk}$ wurde zwar als Traube akzeptiert, enthält jedoch mehrere einzelne Trauben, die während des CC-Verfahrens zu einer Komponente vereint wurden. Die in einer Komponente vereinten Trauben werden daher nicht individuell und die Komponente als lediglich eine einzige Traube gezählt.

Die False Negatives werden wie folgt gezählt:

- Die durch die Gründe 1-4) nicht detektierten Trauben zählen jeweils als ein False Negative.
- Eine Komponente, die mehreren Trauben enthält (Grund 5), wird als ein True Positive gewertet, da richtigerweise mindestens eine Traube gezählt wird. Die Anzahl der in der Komponente enthaltenen Trauben abzüglich Eins, werden hingegen als zusätzliche False Negatives gezählt. Eine Komponente könnte bspw. aus drei tatsächlichen Trauben bestehen und würde als ein True Positive gezählt werden. Die Konstellation würde gleichzeitig zwei False Negatives erzeugen, da die verbleibenden zwei Trauben nicht individuell gezählt werden können, obwohl sie einzelnen Trauben entsprechen.

Die Precision wird durch die Anzahl der detektierten Trauben bestimmt, die keine tatsächliche Traube darstellen ($\mathcal{X}^T$). Sie zählen als False Positives und werden folgendermaßen gezählt:

- Ein einzige Traube kann während des CC-Verfahrens fälschlicherweise in zwei oder mehrere Komponenten unterteilt werden. Hier wird ein True Positive gezählt, da korrekt eine Traube gezählt wird. Die Anzahl der Komponenten, in welche eine einzige Trauben unterteilt wurde, abzüglich Eins werden als entsprechend viele False Positives gezählt. Eine einzige Traube, die (bspw. aufgrund fehlender Punkte) fälschlicherweise in drei Komponenten unterteilt wurde, würde ein True Positive und zwei False Positives erzeugen, da die zwei übrigen Komponenten als zusätzliche, jedoch nicht tatsächlich existierende Trauben gezählt werden.
- Einzelne Komponenten, die keiner Traube entsprechen, zählen als ein einziges False Positive.

Durch diese Art und Weise zu zählen wird berücksichtigt, dass Komponenten gleichzeitig richtige als auch falsche Zählungen verursachen können. So können Komponenten zwar richtigerweise ausschließlich Punkte der Klasse **Traube** enthalten, aber die Zählung der Trauben aufgrund der beschriebenen Situationen (mehrere Trauben in einer Komponente; eine einzige Traube in mehreren Komponenten) verfälschen. Indem eine Komponente gleichzeitig True Positives wie auch False Positives oder False Negatives erzeugen kann, wird diesem Umstand genüge getan.

Referenzwerte für die Beerenanzahl

Referenzwerte für die Beerenanzahl werden durch manuellen Abgleich von $\mathcal{V}^B$, $\mathcal{D}^B$ und $\mathcal{M}^B$ von jeweils zehn

detektierten Trauben pro 5m Ausschnitt mit unterschiedlicher Größe und Beerenanzahl erhalten. Es wird direkt zwischen den Beeren in den Bildern und Ergebnissen für $\mathcal{D}^B$ verglichen. Alle auf den Bildern sichtbaren Beeren zählen zu $\mathcal{V}^B$, alle nicht detektierten Beeren² zählen zu $\mathcal{M}^B$. Kugeln, die keiner tatsächlichen Beeren entsprechen, zählen zu $\mathcal{X}^B$.

Zur Berechnung von Recall und Precision der Beerenzählung werden True Positives, False Negatives und False Positives folgendermaßen gezählt:

- Eine Kugel, die der jeweiligen Beere im Bild entspricht, zählt als ein True Positive.
- Eine Kugel, die mit keiner tatsächlichen Beere korrespondiert, zählt als ein False Positive.
- Eine Beere, die auf den Bildern sichtbar war, aber in keiner Komponente vorhanden ist, zählt als ein False Negative.
- Eine Beere, die in einer Komponente vorhanden ist, aber nicht vom Berryfinder detektiert werden konnte, zählt als ein False Negative.

Referenzwerte für die Beerengröße

Referenzwerte für die Beerengröße ergeben sich aus der manuellen Bestimmung des Durchmessers von 100 Beeren mit einem Messschieber. Zur Evaluierung der Schätzung der Beerengröße in den TS- und MSS-Anbaureihen werden die geschätzten Radien mit den manuell bestimmten Radien der 100 Beeren in Histogrammen verglichen und der Mittelwert von gemessenen und geschätzten Durchmessern berechnet.

4.3.2.3 Bestimmung von Recall und Precision der Beeren- und Traubendetektion für beide Erziehungssysteme und Reifestadien

Tabelle 4.9 listet die gewählten Parameter und Schwellwerte für das CC-Verfahren und den Berryfinder für beide Reifestadien und Erziehungssysteme auf. Die Schwellwerte werden je Erziehungssystem und Reifestadium angepasst.

	TS 75	MSS 75	TS 89	MSS 89
p_K	300	300	400	300
p_{max}	4000	3000	5000	3000
1.findSpheres	$r_N = 2mm$	$r_N = 2mm$	$r_N = 2mm$	$r_N = 2mm$
	$r_{in} = 60\%$	$r_{in} = 60\%$	$r_{in} = 75\%$	$r_{in} = 75\%$
	$\#P^{in} = 15$	$\#P^{in} = 15$	$\#P^{in} = 50$	$\#P^{in} = 50$
2.findSpheres	$r_N = 2mm$	$r_N = 2mm$	$r_N = 4mm$	$r_N = 4mm$
	$r_{in} = 40\%$	$r_{in} = 40\%$	$r_{in} = 60\%$	$r_{in} = 60\%$
	$\#P^{in} = 10$	$\#P^{in} = 10$	$\#P^{in} = 30$	$\#P^{in} = 30$
min_B	5	5	3	2

Tabelle 4.9: Übersicht über die gewählten Schwellwerte des CC-Verfahrens und des Berryfinders für beide Erziehungssysteme und Reifestadien.

Im MSS-Erziehungssystem werden p_K und p_{max} aufgrund der generell kleineren Trauben kleiner als im TS-Erziehungssystem gewählt. Im BBCH75 Reifestadium werden die Schwellwerte des Berryfinders aufgrund

²Die Zählung bezieht sich ausschließlich auf die $\mathcal{M}^B$ detektierter Trauben. Die $\mathcal{M}^B$ nicht detektierter Trauben werden nicht gezählt.

des kleineren Durchmessers der Beeren im Vergleich zum BBCH89 Reifestadium und der dadurch erschwerten Rekonstruktion geringer gewählt, um dennoch die kleineren Beeren detektieren zu können. Aufgrund dieser geringeren zu überwindenden Schwellwerte des Berryfinders wird der Schwellwert $\min_B$ höher als im BBCH89 Reifestadium für beide Erziehungssysteme gewählt, um die Precision der Traubenzählung für das BBCH75 Reifestadium zu erhöhen.

Die Schwellwerte wurden durch empirisches Ausprobieren und visuellen Abgleich bestimmt.

Recall und Precision der Traubendetektion

Für jeden der fünf Abschnitte werden Recall und Precision von $\mathcal{D}^T$ bestimmt, sowie der jeweilige Mittelwert und die Standardabweichung berechnet. Die Anzahl $\mathcal{M}^T$ wird manuell über alle fünf Abschnitte bestimmt und das Verhältnis zu $\mathcal{V}^T$ berechnet. Das Vorgehen wird exemplarisch in Tabelle 4.10 anhand der MSS-Anbaureihe im BBCH75 Reifestadium dargestellt, die übrigen Ergebnisse werden direkt in Tabelle 4.11 zusammengefasst.

Von der TS-Anbaureihe im BBCH75 Reifestadium werden im Durchschnitt 60% aller Trauben detektiert. Von allen Komponenten, die als Traube akzeptiert wurden, sind im Durchschnitt 89% auch tatsächlich Trauben. Beide Werte erreichen eine Standardabweichung von 8%.

	Recall [%]	Precision [%]
Meter 1	65	93
Meter 2	58	100
Meter 3	47	90
Meter 4	64	82
Meter 5	67	80
$\overline{Meter}$	60	89
σ	8	8

Tabelle 4.10: Recall und Precision von $\mathcal{D}^T$ für 5×1 -Meter der TS-Anbaureihe im BBCH75 Reifestadium, sowie ihr Mittelwert und ihre Standardabweichung.

Tabelle 4.11 fasst die Ergebnisse des durchschnittlichen Recalls $\overline{Rc}$ und der durchschnittlichen Precision $\overline{Prc}$ beider Erziehungssysteme und Reifestadien zusammen.

	TS		MSS	
BBCH89	$\overline{Rc} : 71$	$\sigma : 12$	$\overline{Rc} : 68$	$\sigma : 10$
	$\overline{Prc} : 95$	$\sigma : 7$	$\overline{Prc} : 100$	$\sigma : 0$
BBCH75	$\overline{Rc} : 60$	$\sigma : 8$	$\overline{Rc} : 64$	$\sigma : 7$
	$\overline{Prc} : 89$	$\sigma : 8$	$\overline{Prc} : 93$	$\sigma : 4$

Tabelle 4.11: Durchschnittlicher Recall $\overline{Rc}$ und Precision $\overline{Prc}$ von $\mathcal{D}^T$ aus 5×1 -Meter für beide Reifestadien und Erziehungssysteme.

Bei der Detektion der Trauben im BBCH89 Reifestadium werden für beide Reifestadien jeweils ähnliche Werte für den Recall in beiden Erziehungssystemen erreicht. Bei der TS- bzw. der MSS-Anbaureihe werden durchschnittlich 71% bzw. 68% der Trauben detektiert, bei ähnlicher Standardabweichung von 12% bzw. 10%.

Für das Reifestadium BBCH75 ergeben sich ebenfalls ähnliche Werte für den Recall. Der Recall ist im Vergleich zum BBCH89 Reifestadium mit 60% für die TS-Anbaureihe und 64% für die MSS-Anbaureihe um rund 10%-Punkte geringer, die Standardabweichung ist etwa gleich hoch bei 8% bzw. 7%.

Tabelle 4.12 listet zusammenfassend $\mathcal{V}^T$, $\mathcal{M}^T$ sowie deren Verhältnis $\Delta = \mathcal{M}^T/\mathcal{V}^T[\%]$ über alle fünf Meter beider Erziehungssysteme und Reifestadien auf. Das Verhältnis $\Delta[\%]$ wird nach Gründen für die Nicht-Detektion aufgeschlüsselt.

	(I)	(II)	(III)	(1)	(2)	(3)	(4)	(5)
	$\mathcal{V}^T$	$\mathcal{M}^T$	$\Delta[\%]$	Nicht rekonstruiert	Falsch klassifiziert	Punktzahl $< p_K$	Beerenanzahl $< min_B$	Nicht trennbar
				[%]	[%]	[%]	[%]	[%]
TS 89	80	27	34	11	0	0	14	9
MSS 89	73	24	33	19	0	1	12	0
TS 75	81	28	34	10	0	10	7	7
MSS 75	123	44	36	14	0	11	9	2

Tabelle 4.12: Spalten I-III: $\mathcal{V}^T$, $\mathcal{M}^T$ und ihr Verhältnis $\Delta = \mathcal{M}^T/\mathcal{V}^T$ über alle fünf Meter beider Erziehungssysteme und Reifestadien. Spalten 1-5 schlüsseln $\Delta[\%]$ geordnet nach Gründen für die Nicht-Detektion auf.

Das Verhältnis $\Delta[\%]$ von $\mathcal{M}^T$ an $\mathcal{V}^T$ liegt über die fünf Meter Abschnitte bei ungefähr einem Drittel (33 – 36%) für beide Erziehungssysteme und Reifestadien. Die Gründe für die False Negatives gestalten sich anteilig zwar unterschiedlich an $\Delta[\%]$, korrespondieren jedoch mit Reifestadium und Erziehungssystem:

1. Nicht rekonstruiert: Im MSS-Erziehungssystem werden, unabhängig vom Reifestadium, ein größerer Anteil der Trauben durch Pix4DMapper nicht rekonstruiert. Viele dieser Trauben liegen tief im Inneren des Blattwerks oder bereits auf der gegenüberliegenden Seite der Anbaureihe. Zur Rekonstruktion eines Objektpunktes muss dieser mindestens dreimal in verschiedenen Bildern beobachtet worden sein. Der Verdeckungsgrad verhinderte in diesen Fällen eine ausreichend häufige Beobachtung. Im TS-Erziehungssystem liegen die Trauben offener und können daher öfter für die Rekonstruktion beobachtet werden. Eine höhere Anzahl an Perspektiven und die Erfassung der gegenüberliegenden Seite der Anbaureihe könnte den Verdeckungsgrad weiter verringern. Dadurch würde die Wahrscheinlichkeit steigen, dass tiefer liegende Trauben häufiger erfasst und rekonstruiert werden können.
2. Falsch klassifiziert: Unabhängig vom Erziehungssystem oder Reifestadium wurden in keinem Fall alle 3D-Punkte einer Traube vollständig falsch klassifiziert. Lediglich kleinere Anteil der rekonstruierten Trauben wurden falsch klassifiziert, diese Punkte konstituierten jedoch niemals eine vollständige Traube. Der Großteil der 3D-Punkte der Trauben konnte für die weiteren Detektionsschritte genutzt werden.
3. Punktzahl $< p_K$: Im BBCH75 Reifestadium werden im Vergleich zum BBCH89 Reifestadium, unabhängig vom Erziehungssystem, deutlich mehr Trauben entfernt, weil sie die nötige Punktzahl p_K nicht erreichen. Trauben sind in diesem Reifestadium kleiner und bestehen aus weniger 3D-Punkten als im BBCH89 Reifestadium. Werden Trauben nur teilweise rekonstruiert oder korrekt klassifiziert, bestehen sie teilweise aus einer Anzahl von 3D-Punkten $< p_K$. Die Punktzahl von Komponenten, die keine Trauben enthalten und durch p_K entfernt werden sollen, bleibt jedoch in beiden Reifestadien ungefähr identisch. Ein geringerer Schwellwert für p_K würde zwar die Anzahl fälschlich entfernter Trauben reduzieren, jedoch gleichzeitig in einer erhöhten Anzahl von akzeptierten Komponenten resultieren, die keiner Traube entsprechen (False Positives). Um diese Komponenten zu entfernen, wird p_K daher entsprechend groß und identisch zum BBCH89 Reifestadium gewählt. In der Folge werden im BBCH75 Reifestadium vermehrt fälschlicherweise Komponenten mit einer Punktzahl $< p_K$ entfernt, die Trauben enthalten. Abbildung 4.34 zeigt exemplarisch anhand der TS-Anbaureihe im BBCH89 Reifestadium den Effekt des CC-Verfahrens und des Schwellwerts p_K auf die als **Traube** klassifizierten Punkte.
4. Beerenanzahl $< min_B$: Die Schwellwertdiskriminierung nach der minimalen Beerenanzahl min_B entfernt Traubenkandidaten, die keine Trauben enthalten. Auffallend ist, dass im Reifestadium BBCH89 mit $min_B = 2 - 3$ mehr tatsächliche Trauben fälschlicherweise entfernt werden, obwohl im BBCH75 Reifestadium der zu überwindende Schwellwert mit $min_B = 5$ höher angesetzt wurde. Dabei spielen höchstwahrscheinlich die gesetzten Schwellwerte der Funktion *findSpheres* des Berryfinders eine Rolle.

Im BBCH75 Reifestadium sind die Schwellwerte in *findSpheres* in Vergleich zum BBCH89 Reifestadium niedriger angesetzt. Kugelschätzungen werden daher eher als Beeren akzeptiert und mit höherer Wahrscheinlichkeit eine Anzahl Beeren $> min_B$ detektiert. Dadurch werden weniger Trauben fälschlicherweise entfernt.

5. Nicht trennbar: Das CC-Verfahren löscht den Großteil der falsch klassifizierten Punkte und ist in der Lage, auch dicht benachbarte Trauben zu unterscheiden, die $1 - 2mm$ auseinanderliegen. Die Methode versagt jedoch teilweise bei den dicht benachbarten Trauben der TS-Erziehung. In beiden Reifestadien werden deutlich häufiger als im MSS-Erziehungssystem mehrere Trauben zu einer einzigen Traube zusammen gefasst, weil ihre Distanz zu gering für eine räumliche Unterscheidung über das CC-Verfahren ist.

In MSS-Anbaureihen wachsen Trauben räumlich weiter auseinander und nur selten dicht zusammen. Sie sind daher mit dem CC-Verfahren räumlich leichter zu trennen, so dass Zusammenfassungen mehrerer Trauben in einer Komponente in beiden Reifestadien des MSS-Erziehungssystems nur selten vorkommen. Abbildung 4.35 zeigt die Verteilung der Trauben im Blattwerk für das BBCH89 Reifestadium in beiden Erziehungssystemen.

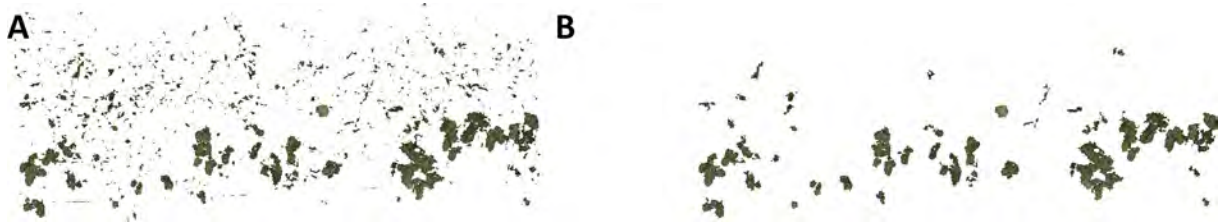


Abbildung 4.34: **A:** 3D-Punkte, die der Klasse **Traube** zugewiesen wurden. Kleine, unstrukturierte Punktreionen sind zum Großteil Fehlklassifikationen. **B:** Das CC-Verfahren entfernt alle Komponenten mit einer Punktzahl $< p_K$. Dadurch wird der Großteil der falsch klassifizierten Punkte entfernt, während der Großteil der korrekt klassifizierten Punkte erhalten bleibt, aus Rose et al. [2016].

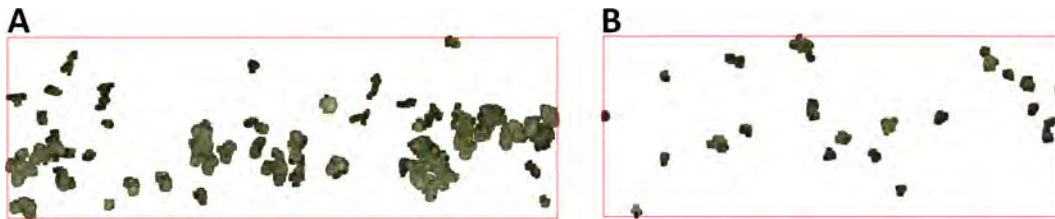


Abbildung 4.35: **A:** Typische Verteilung der Traubenkandidaten in der TS-Anbaureihe. Trauben sind größer und wachsen deutlich dichter zusammen. Ihre räumliche Trennung durch das CC-Verfahren versagt daher öfter als im MSS-Erziehungssystem. **B:** Typische Verteilung der Trauben in der MSS-Anbaureihe. Trauben sind deutlich kleiner und liegen räumlich weiter auseinander. Im CC-Verfahren können diese Punktregionen problemlos zu individuellen Komponenten zusammengefasst werden.

Die Schwellwerte p_K und min_B bestimmen auch die Precision. Trauben werden durch die Schwellwerte zwar u. U. fälschlicherweise entfernt, jedoch gelingt durch die gesetzten Schwellwerte die Detektion der Trauben mit einer hohen Precision von mind. 89% und max. 100% für beide Reifestadien und Erziehungssysteme.

Die Precision im BBCH89 Reifestadium ist für beide Erziehungssystem mit über 95% für die TS-Anbaureihe und 100% für die MSS-Anbaureihe sehr hoch, bei einer geringen Standardabweichung von max. 7%. In diesem Reifestadium werden die meisten Trauben fälschlicherweise durch min_B entfernt (vgl. Tabelle 4.12), was darauf hinweist, dass dieser Schwellwert ebenfalls die meisten Traubenkandidaten filtert, die keine Traube darstellen (False Positives). Aufgrund der hohen Precision bei einer geringen Standardabweichung im BBCH89 Reifestadium kann festgestellt werden, dass in diesem Reifestadium beinahe sämtliche als Traube akzeptierte Komponenten auch tatsächlichen Trauben entsprechen. Abbildung 4.36 illustriert den Einfluss des Schwellwertes für die minimale Beerenanzahl min_B auf die Traubenzählung.

Die mittlere Precision ist im BBCH75 Reifestadium vergleichsweise etwas niedriger, jedoch mit 89% für die TS-Anbaureihe bzw. 93% für die MSS-Anbaureihe immer noch hoch, mit einer Standardabweichung von 8% bzw. 4%. In diesem Reifestadium werden die meisten tatsächlichen Trauben über p_K entfernt, dieser Schwellwert filtert somit die meisten Traubenkandidaten, die keine Trauben darstellen. Der Schwellwert min_B spielt eine vergleichsweise geringere Rolle, was aus den niedrigeren Schwellwerten für *findSpheres* resultiert. Da im BBCH75 Reifestadium aufgrund der niedrigen Schwellwerte für *findSpheres* durchschnittlich mehr Kugelschätzungen als gültige Beeren akzeptiert werden, steigt auch die Anzahl der dadurch fälschlicherweise als Traube akzeptierten Traubenkandidaten, die tatsächlich keine Trauben abbilden. Dies resultiert in einer geringeren Precision für das BBCH75 Reifestadium im Vergleich zum BBCH89 Reifestadium.

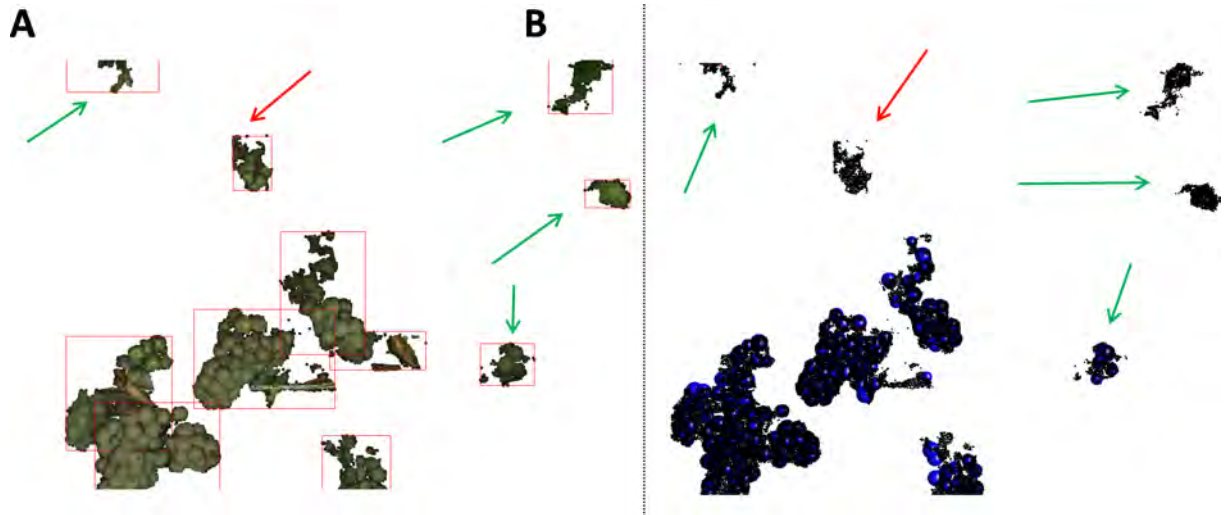


Abbildung 4.36: **A:** Beispielhafte Traubenkandidaten $\in P^{T_k}$ in ihrer rot umrahmten Komponente. Kleinere Punktregionen sind entweder Regionen falsch klassifizierter Punkte oder Trauben, die nur teilweise vorhanden sind. **B:** Detektierte Beeren und als tatsächliche Traube akzeptierte Traubenkandidaten. Traubenkandidaten mit dargestellten Kugeln sind akzeptierte Traubenkandidaten. Die Schwellwertdiskriminierung nach der Mindestanzahl an Beeren min_B während des Berryfinders entfernt den Großteil von Komponenten mit falsch klassifizierter Punkten und detektiert korrekt kleinere Traubenregionen (grüne Pfeile). Einige kleinere Trauben werden dabei dennoch unter Umständen verworfen, weil bei ihnen nicht genügend Beeren enthalten waren (roter Pfeil).

Recall und Precision der Beerendetektion

Für jeweils 10 Trauben unterschiedlicher Größe jedes Erziehungssystems und Reifestadiums werden Recall und Precision von $\mathcal{D}^B$ bezogen auf $\mathcal{V}^B$ bestimmt. Aus den 10 Werten werden jeweils Mittelwert und Standardabweichung berechnet. Die Ergebnisse für die TS-Anbaureihe im BBCH75 Reifestadium werden beispielhaft in Tabelle 4.13 vollständig dargestellt, alle weiteren Ergebnisse direkt tabellarisch in Tabelle 4.14 zusammengefasst.

Durchschnittlich werden für die TS-Anbaureihe im BBCH75 Reifestadium 76% aller sichtbaren Beeren detektiert und durchschnittlich 92% aller detektierten Beeren entsprechen tatsächlichen Beeren. Die Standardabweichung beträgt 14% bzw. 3% (Tabelle 4.13).

Der durchschnittliche Recall der Beerendetektion ist für beide Reifestadien und Erziehungssysteme ähnlich, mit ebenfalls ähnlichen Standardabweichungen. Der durchschnittliche Recall liegt bei 72 – 78%, mit einer Standardabweichung bei 11 – 14% (Tabelle 4.14). Beeren werden aus mehreren Gründen nicht detektiert:

1. Beeren einer Traube werden nicht rekonstruiert, da sie nicht häufig genug beobachtet werden können. Diese Beeren liegen typischerweise tiefer in der Traube und sind häufig von vordergründig liegenden Beeren oder Blättern verdeckt.

2. Beeren werden nicht als zur Klasse **Traube** gehörig klassifiziert. Diese Beeren liegen häufig in der Nähe von Ästen oder Blättern. Dies erschwert ihre Klassifikation, da die lokale Geometrie von der typischen Kugelform abweicht.
3. Die Beere ist in der Punktwolke vorhanden, kann jedoch nicht vom Berryfinder detektiert werden. Dies geschieht, wenn ihre Rekonstruktion so sehr von der Kugelgeometrie abweicht, dass die Kugelschätzung mit den gesetzten Schwellwerten des Berryfinders nicht mehr als Beere akzeptiert werden kann.

	$\mathcal{V}^B$	Recall [%]	Precision [%]
Traube 1	71	76	92
Traube 2	115	79	89
Traube 3	45	91	89
Traube 4	59	97	95
Traube 5	76	70	93
Traube 6	36	44	94
Traube 7	32	69	88
Traube 8	31	84	93
Traube 9	43	72	91
Traube 10	44	75	94
$\overline{\text{Traube}}$		76	92
σ		14	3

Tabelle 4.13: $\mathcal{V}^B$, Recall und Precision für $\mathcal{D}^B$ für 10 Trauben unterschiedlicher Größe für die TS-Anbaureihe im BBCH75 Reifestadium, sowie der jeweilige Mittelwert und die Standardabweichung.

	TS	MSS
BBCH89	$\overline{Rc} : 78 \sigma : 13$ $\overline{Prc} : 98 \sigma : 3$	$\overline{Rc} : 77 \sigma : 11$ $\overline{Prc} : 98 \sigma : 3$
BBCH75	$\overline{Rc} : 76 \sigma : 14$ $\overline{Prc} : 92 \sigma : 3$	$\overline{Rc} : 72 \sigma : 11$ $\overline{Prc} : 88 \sigma : 11$

Tabelle 4.14: Durchschnittlicher Recall $\overline{Rc}$ und Precision $\overline{Prc}$ für die Zählung von $\mathcal{D}^B$ aus 5×1 -Meter für beide Reifestadien und Erziehungssysteme.

Die Precision liegt im BBCH89 Reifestadium für beide Erziehungssysteme bei 98%, bei einer geringen Standardabweichung von 3%. Die Beeren sind in diesem Reifestadium größer und daher leichter zu rekonstruieren. Da zudem kaum andere Bestandteile der Anbaureihen wie Äste oder Blätter in den Traubenkandidaten enthalten sind, entsprechen beinahe sämtliche der Kugelschätzungen tatsächlichen Beeren. Die Precision verringert sich im BBCH75 Reifestadium auf 92% für die TS-Anbaureihe und 88% für die MSS-Anbaureihe. Die Standardabweichung der Precision der TS-Anbaureihe beträgt 3% und 11% für die MSS-Anbaureihe. Das Ergebnis der MSS-Anbaureihe resultiert höchstwahrscheinlich aus der geringeren Kompaktheit der Trauben im MSS-Erziehungssystem zu diesem Reifestadium. Das Stielgerüst liegt offener und ist teilweise gut sichtbar. Ein Teil des Stielgerüsts der Traube wird daher ebenfalls rekonstruiert und korrekt als Punkte der Klasse **Traube** klassifiziert. Durch die in diesem Reifestadium geringen Schwellwerte der Funktion *findSpheres* des Berryfinders werden jedoch wahrscheinlicher Kugeln als Beeren akzeptiert, die eigentlich zum Stielgerüst gehören. Diese False Positives führen zu einer geringeren Precision mit vergleichsweise höherer Standardabweichung.

Abbildungen 4.37-4.40 zeigen exemplarisch mehrere Trauben der TS- bzw. der MSS-Anbaureihe in beiden Reifestadien und die vom Berryfinder geschätzten Beeren. Gut erkennbar ist die unterschiedliche Größe der

Trauben und die Anzahl der Beeren der verschiedenen Erziehungssysteme und Reifestadien. Der Großteil der Beeren wurde detektiert und korrespondiert mit den jeweiligen tatsächlichen Beeren.

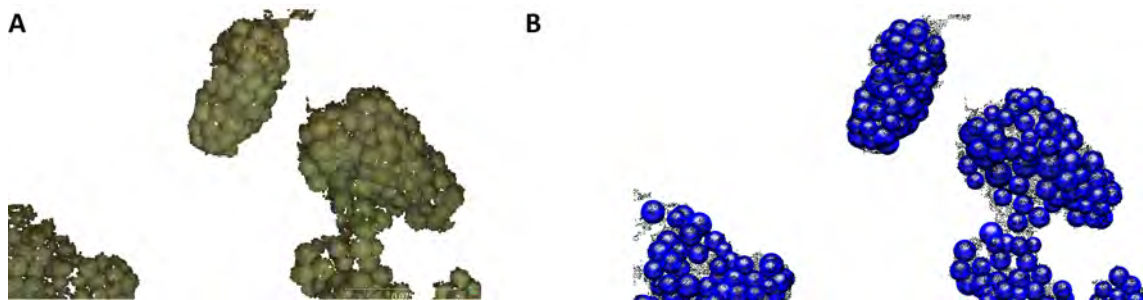


Abbildung 4.37: **A:** Punktwolken mehrerer Traubenkandidaten der TS-Anbaureihe im BBCH89 Reifestadium. **B:** Detektierte Beeren in den Traubenkandidaten, aus Rose et al. [2016].

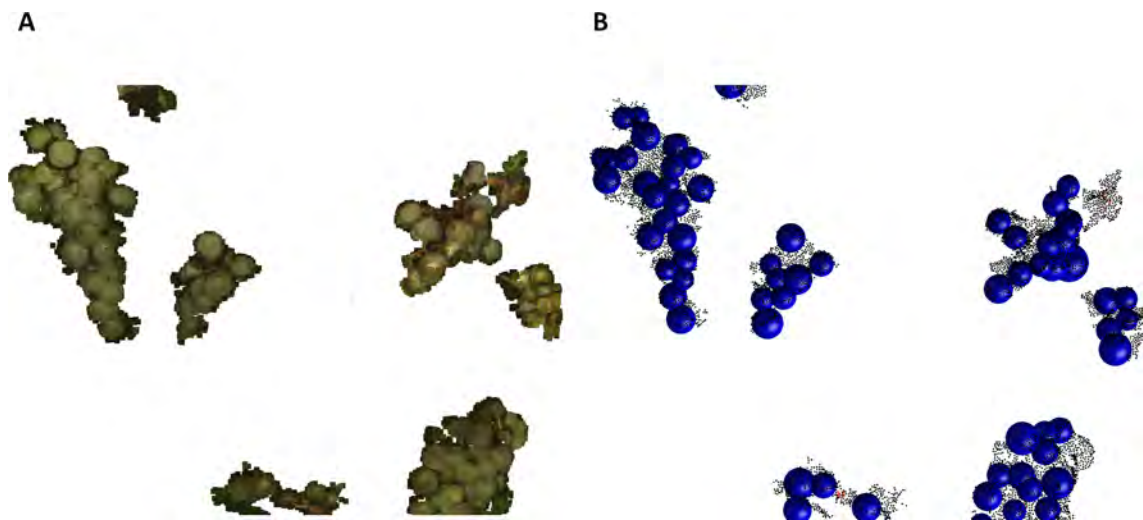


Abbildung 4.38: **A:** Punktwolken mehrerer Traubenkandidaten der MSS-Anbaureihe im BBCH89 Reifestadium. **B:** Detektierte Beeren in den Traubenkandidaten.

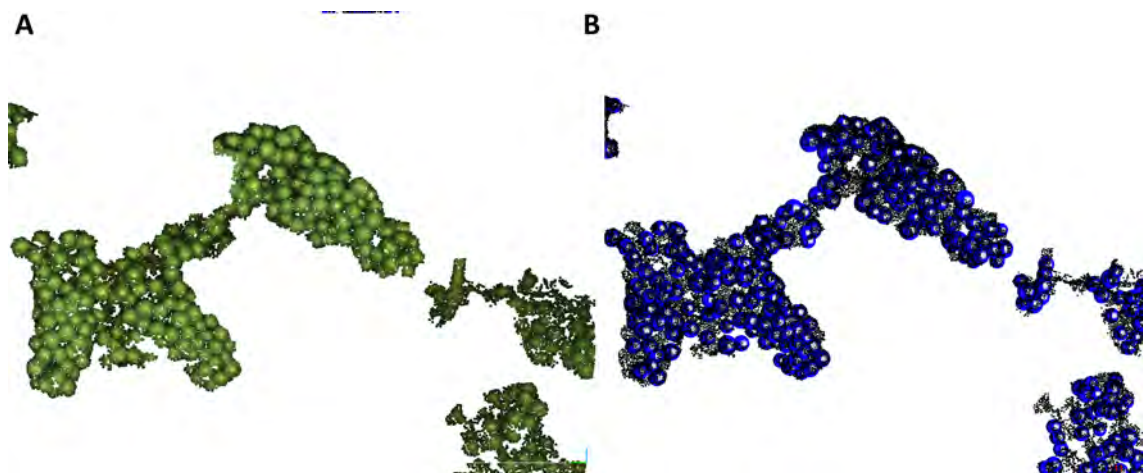


Abbildung 4.39: **A:** Punktwolken mehrerer Traubenkandidaten der TS-Anbaureihe im BBCH75 Reifestadium. **B:** Detektierte Beeren in den Traubenkandidaten.

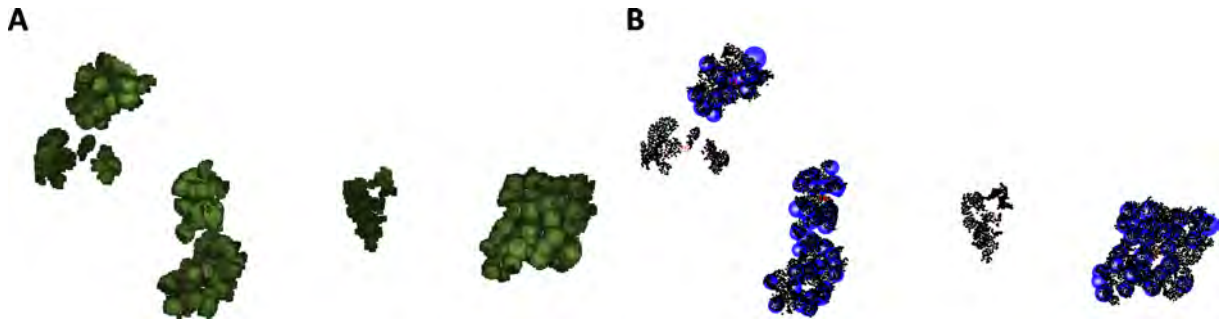


Abbildung 4.40: **A:** Punktwolken mehrerer Traubenkandidaten der MSS-Anbaureihe im BBCH75 Reifestadium. **B:** Detektierte Beeren in den Traubenkandidaten.

4.3.2.4 Vergleich von geschätzter und gemessener Beerengröße

Tabelle 4.15 listet die Gesamtanzahlen $\mathcal{D}^B$ der fünf Abschnitte beider Erziehungssysteme und Reifestadien, sowie die mittlere geschätzte und manuell gemessene Beerengröße mit ihren empirischen Standardabweichungen auf. Die Gesamtanzahl der Beeren ergibt sich durch Summierung der Beerenanzahl pro detektierter Traube. Die geschätzte durchschnittliche Beerengröße entspricht zu $1 - 2,5\text{mm}$ den durchschnittlichen Beerengrößen der Referenzmessungen für beide Erziehungssysteme und Reifestadien. Dabei ist die Schätzung, unabhängig von Reifestadium oder Erziehungssystem, stets größer als die tatsächliche Messung.

	$\mathcal{D}^B$	Geschätzter mittlerer Beeren Durchmesser (mm)	Gemessener mittlerer Beeren Durchmesser (mm)
TS 89	1437	$\overline{13,7} \pm 1,8$	$\overline{11,8} \pm 1,4$
MSS 89	567	$\overline{12,9} \pm 1,7$	$\overline{10,4} \pm 1,6$
TS 75	1593	$\overline{10,1} \pm 1,2$	$\overline{9,0} \pm 0,6$
MSS 75	2030	$\overline{9,9} \pm 1,2$	$\overline{7,8} \pm 0,6$

Tabelle 4.15: Detektierte Gesamtanzahl $\mathcal{D}^B$ aus allen akzeptierten Trauben, mittlere geschätzte und manuell gemessene Beerengröße mit jeweiliger Standardabweichung für 5-Meter beider Erziehungssysteme und Reifestadien.

T-Test von geschätzter und gemessener Beerengröße

Um zu überprüfen, ob sich die Mittelwerte von Schätzung und manueller Messung signifikant unterscheiden, wird jeweils für Reifestadium und Erziehungssystem ein zweiseitiger T-Test für unabhängige Stichproben durchgeführt [Benning, 2011]. Es wird angenommen, dass die Standardabweichungen beider Stichproben unbekannt, jedoch gleich ist und die Beerengröße der Normalverteilung folgt, da sie aus der gleichen Anbaureihe stammen.

Beim T-Test wird die t-Verteilung genutzt. Die Verteilungen der geschätzten und gemessenen Beerengrößen sind in den Histogrammen in Abbildung 4.41 und 4.42 dargestellt. Es wird angenommen, dass die geschätzte durchschnittliche Beerengröße der gemessenen durchschnittlichen Beerengröße entspricht. Die Nullhypothese lautet somit:

$$H_0 : \mu_1 = \mu_2, \quad (4.16)$$

wobei μ_1 und μ_2 den Mittelwert der Schätzung bzw. der Messung bezeichnen. Die Alternativhypothese lautet:

$$H_1 : \mu_1 \neq \mu_2. \quad (4.17)$$

Es wird ein Signifikanzniveau $\alpha = 5\%$ gewählt. Es ergaben sich folgende Testentscheidungen (Tabelle 4.16):

	H_0 :
TS 89	Wird abgelehnt, H_1 wird angenommen.
MSS 89	Wird abgelehnt, H_1 wird angenommen.
TS 75	Wird abgelehnt, H_1 wird angenommen.
MSS 75	Wird abgelehnt, H_1 wird angenommen.

Tabelle 4.16: Ergebnisse des zweiseitigen T-Tests für beide Reifestadien und Erziehungssysteme.

Die Nullhypothese wird auf dem Signifikanzniveau von 5% für alle Anbaureihen abgelehnt, die geschätzte Beerengröße weicht signifikant von der gemessenen Beerengröße ab. Die geschätzte Beerengröße entspricht damit **nicht** der tatsächlichen Beerengröße.

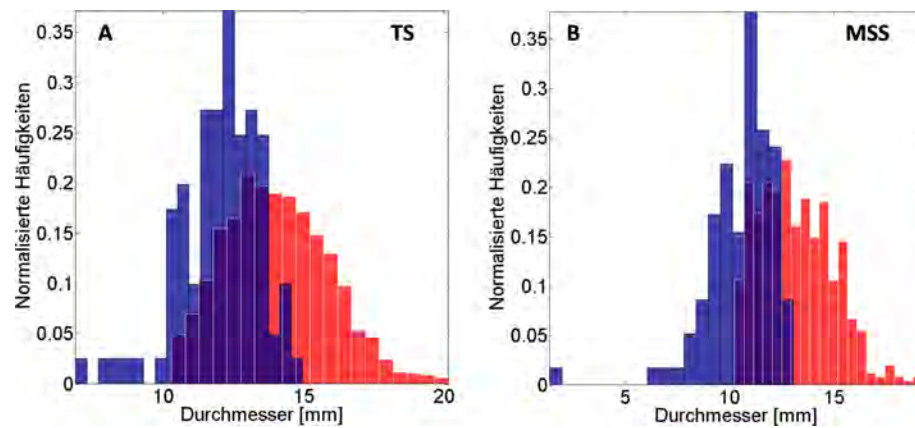


Abbildung 4.41: Normalisierte Histogramme für die Schätzung der Beerengröße im BBCH89 Reifestadium. **A:** TS-Anbaureihe. **B:** MSS-Anbaureihe. **Rot:** Schätzung durch den Berryfinder. **Blau:** Manuelle Messung. **Violett:** Überlappungsbereich beider Histogramme.

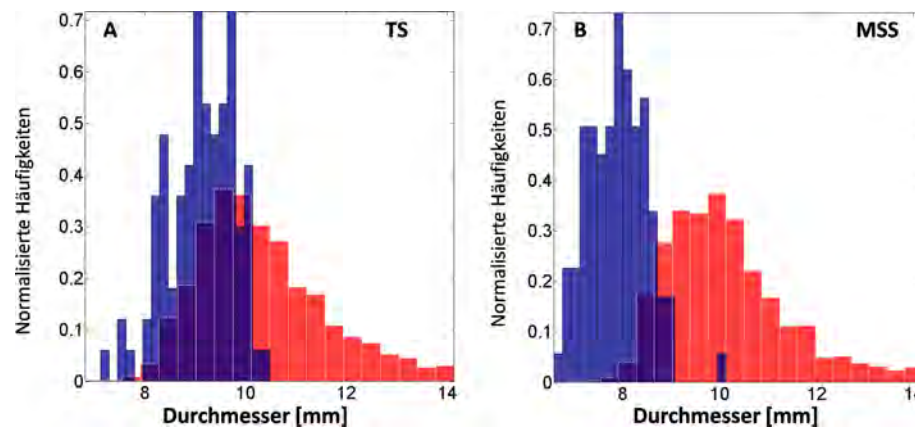


Abbildung 4.42: Normalisierte Histogramme für die Schätzung der Beerengröße im BBCH75 Reifestadium. **A:** TS-Anbaureihe. **B:** MSS-Anbaureihe. **Rot:** Schätzung durch den Berryfinder. **Blau:** Manuelle Messung. **Violett:** Überlappungsbereich beider Histogramme.

Die Abweichung zwischen Schätzung und Messung kann mehrere Gründe besitzen:

- $\mathcal{D}^B$ ist immer deutlich größer als die nur 100 Beeren umfassende Stichprobe der Referenzmessungen, was die Unterschiede in der Verteilung erklären kann. Nicht ausgeschlossen werden kann auch die

Möglichkeit, dass ein Anwender unbewusst vermehrt kleinere Beeren in die Referenzmessungen mit einbezieht, um alle Beerengrößen zu berücksichtigen. Diese kleineren Beeren werden jedoch u. U. gar nicht rekonstruiert, weil sie im Inneren der Traube oder an schlecht einsehbaren Stellen liegen.

- Durch die Auflagepunkte der Zange und natürliche Abweichungen der Beere von einer idealen Kugel ergibt sich eine Unsicherheit der Messungen mit dem Messschieber. So ist zu beachten, dass die Kugelschätzung auf dem Teil der Beere stattfindet, welcher der Kamera zugewendet ist. Die Messung mit einem Messschieber geschieht dagegen primär an seitlich zu diesem Segment gelegenen Stellen der Beere, welche einen leicht abweichenden Radius bilden können. Ein leichtes Zusammendrücken der Beeren durch den Messschieber kann ebenfalls nicht ausgeschlossen werden.
- Unterschiede können auch aus Abweichungen der rekonstruierten Beeren von ihrer tatsächlichen Geometrie entstehen (vgl. Kapitel 4.3.1.2). Wurde die Krümmung der Beeren etwas abgeflachter rekonstruiert und/oder durch die Oberflächenglättung reduziert, werden tendenziell zu große Kugeln geschätzt. Die Rekonstruktion der Beeren wird insbesondere im BBCH75 Reifestadium durch ihre geringe Größe und die Reflexion auf der Beerenmitte erschwert, weshalb Beeren in diesem Reifestadium häufig abgeflacht rekonstruiert werden (s. Abbildung 4.9). Aus diesem Umstand lässt sich auch die vergleichsweise größere Standardabweichung der Schätzungen im Vergleich zu den manuellen Messungen erklären. Diese betrug im BBCH75 Reifestadium für die geschätzte Beerengröße $1,2\text{mm}$, für die gemessene Beerengröße hingegen nur $0,6\text{mm}$.
- Die Schwellwerte des Berryfinders sind zwar optimal für die Detektion der Beeren, jedoch nicht für die Schätzung der Beerengröße. Weicht die Rekonstruktion der Beeren zu sehr von der Kugelform ab, kann die Detektion der Beeren dennoch mit guten Ergebnissen erfolgen, wenn die Schwellwerte von findSpheres entsprechend gewählt werden. Die dann geschätzte Beerengröße entspricht jedoch nicht zwangsläufig der tatsächlichen Beerengröße.

4.4 Zwischenfazit: Voruntersuchungen mit dem PHENObot

Von den Trauben werden durchschnittlich 60 – 71% detektiert. Über alle fünf Meter ergab sich für alle vier Anbaureihen ein ähnliches Verhältnis von $\mathcal{M}^T$ an $\mathcal{V}^T$ von etwa einem Drittel (Tabelle 4.12, 33 – 36%). Es lässt sich daher konstatieren, dass die sensorgestützte Detektion der Ertragsparameter nach der hier vorgestellten Methode zwar nicht alle sichtbaren Trauben detektiert, die Auswertemethode über mehrere Meter von Weinreben summiert für beide Erziehungssysteme und Reifestadien jedoch weitestgehend konstante Ergebnisse erzeugt (vgl. Nuske et al. [2011], wo ebenfalls ein gleichbleibender Anteil der Detektion beobachtet wird). Die mittlere Precision der Traubendetektion ist mit Werten zwischen 89 – 100% sehr hoch, so dass konstatiert werden kann, dass beinahe sämtliche detektierte Trauben tatsächlichen Trauben entsprechen.

Der mittlere Recall der Beerendetektion ist, unabhängig von Erziehungssystem und Reifestadium, mit Werten zwischen 72 – 78% weitestgehend konstant bei ähnlichen Standardabweichungen. Auch die mittlere Precision ist mit über 88 – 98% sehr hoch. Zusammenfassend lässt sich gleiches wie für die Traubendetektion konstatieren. Es werden nicht alle sichtbaren Beeren detektiert, jedoch entsprechen aufgrund der hohen Precision beinahe sämtliche detektierten Beeren tatsächlichen Beeren.

Die Bestimmung der durchschnittlichen Beerengröße entspricht bis auf 1 – $2,5\text{mm}$ der durchschnittlichen Beerengröße aus Referenzmessungen, was eine signifikante Abweichung darstellt. Die geschätzte Beerengröße entspricht somit nicht der tatsächlichen Beerengröße. Das Gewicht einer Beere sollte daher nicht über ihre geschätzte Größe bestimmt werden.

Die vorgestellte Auswertemethode für die Detektion der sichtbaren Trauben und Beeren ist für beide Erziehungssysteme und Reifestadien geeignet. Die gesammelten Erfahrungen mit dem PHENObot werden im nächsten Kapitel auf die Konstruktion der neuen Sensorplattform Phenoliner übertragen und mit ihr weitere Untersuchungen durchgeführt. Der Fokus wird darauf liegen, zusätzliche Referenzwerte zu erheben, die Anzahl *aller* Beeren und Trauben (i. e. die Ertragsparameter) zu bestimmen und aus ihnen den Ertrag zu schätzen.

5. Praxisbeispiel 2: Ertragsschätzung mit dem Phenoliner

Aus dem vorherigen Kapitel 4.1 wurde ersichtlich, dass der PHENObot folgende in Kapitel 1 formulierte Kriterien an die Datenerhebung nicht erfüllt:

1. Der PHENObot ist nicht für die Datenerhebung der geforderten drei Kilometer Anbaureihen an einem Arbeitstag geeignet. Nur eine RGB-Kamera ist vorhanden. Insgesamt sind daher drei Durchfahrten durch die Weinreihe in verschiedenen Kamerahöhen notwendig, was pro Anbaureihe rund 15 Minuten in Anspruch nimmt.
2. Eine standardisierte, unabhängige Datenerhebung ist nicht möglich, da Sensorik und Weinreben nicht von äußeren Lichtverhältnissen abgeschirmt sind. Homogene Lichtverhältnisse können nur nach Einbruch der Dunkelheit erreicht werden, was die Nutzungsmöglichkeiten des PHENObots einschränkt.

Der Phenoliner wurde mit der Intention entworfen, die oben genannten Schwächen des PHENObots zu kompensieren und eine Datenerhebung im Hochdurchsatz unabhängig von der Tageszeit zu ermöglichen. Grundlegende Unterschiede sind die Verwendung eines vertikal angeordneten Multi-Kamerasystems aus drei Kameras und eines modifizierten Vollerntetraktors, in dem das MKS sowie weitere Sensoren abgeschirmt vor äußeren Einflüssen montiert sind.

Im Folgenden wird zunächst der Phenoliner detailliert beschrieben. Darauf folgt eine Schilderung der Datenerhebung, welche in Bezug zur Datenerhebung mit dem PHENObot gesetzt wird. Dabei wird auf die Beleuchtung, Bildqualität, die Genauigkeit der Punktwolke und Schwierigkeiten bei der Datenerhebung eingegangen.

Im Anschluss erfolgt die Schilderung der Datenklassifikation der mit dem Phenoliner erhobenen Daten und ein Vergleich zur Datenklassifikation aus dem vorherigen Kapitel.

Zuletzt wird die Schätzung der Gesamtanzahl der Ertragsparameter aus den mit dem Phenoliner erhobenen Daten beschrieben und eine darauf basierende Schätzung des Ertrages vorgenommen. Die geschätzten Erträge werden zuletzt mit manuell gewogenen Referenzerträgen verglichen.

5.1 Datenerhebung

5.1.1 Der Phenoliner und seine Komponenten

Die Funktionalität des Phenoliners wird über drei Komponenten erreicht: 1) einem modifizierten Vollernter-Traktor als Sensorplattform 2) einem Multi-Kamera-System bestehend aus drei vertikal übereinander angeordneten Kameras und 3) dem IggGeotagger 2.0, welcher die Aufnahme der Daten und die Georeferenzierung der Bilder steuert. Die drei Komponenten werden im Weiteren erläutert.

5.1.1.1 Sensorplattform

Die Basis des Phenoliners bildet ein modifizierter ERO-Grapeliner SF200 Traktor. Ein 6-Zylinder Motor mit 157 PS erlaubt Geschwindigkeiten bis zu 40km/h . Alle für die Ernte vorgesehenen Teile im Inneren des Phenoliners wurden entfernt, so dass in der inneren Mitte des Phenoliners ein durchgängiger Tunnel von $0,86\text{m}$ Breite und einer variablen Höhe von $2 - 2,8\text{m}$ entlang der Längsachse des Fahrzeugs entsteht. In dem Tunnel wird das bereits erwähnte Multi-Kamerasystem (MKS) geschützt vor Umwelteinflüssen auf der

rechten Längsseite des Tunnels montiert (s. Kapitel 5.1.1.2). Die Höhe des Tunnels über dem Boden lässt sich über ein Hydrauliksystem direkt aus dem Führerhäuschen optimal für die Beobachtung der Weinreben durch das MKS einstellen.

Die dem MKS gegenüberliegende, linke Tunnelseite wird als Hintergrund bezeichnet. Die Hintergrundfarbe wird durchgehend homogen und in einer für Weinreben untypischen Farbe gehalten. Die homogene Farbgestaltung wird im späteren Verlauf bei der Vorverarbeitung der Punktwolken zur Trennung von Weinreben und Hintergrund benötigt. Um möglichst einheitliche Beleuchtungsverhältnisse unabhängig vom Tageslicht zu erreichen, wurden der hintere und vordere Tunnelausgang, sowie der Großteil aller weiteren Öffnungen verhangen. Für die hier geschilderten Untersuchungen wird ein schweres weißes Tuch zur Gestaltung des Hintergrundes verwendet.

Auf dem oberen Fahrzeugrahmen wurde ein RTK (real-time-kinematic)-GPS System (Trimble® SPS852, Sunnyvale, Kalifornien, USA) installiert, welches die GPS-Position der Antenne mit 2cm Genauigkeit bestimmt. Abschattungen und Mehrwegeeffekte durch die Anbaureihen können durch die Höhe der Antenne von über 4m ausgeschlossen werden. Ein Zwei-Achsen Inklinometer (DOG2 MEMS-Serie USB Rev.1.; TE Connectivity Sensors, Dortmund, Deutschland) für die Berücksichtigung der Rotation des Phenoliners entlang seiner Längsachse (Roll) und Querachse (Pitch) wurde nahe des MKS parallel zur Längsachse des Fahrzeugs installiert. Der Phenoliner ist mit einem PC (Intel Core i7-860 mit $2,8\text{ GHz}$, 4 GB-DDR RAM , $2 \times 480\text{ GB SSD}$ Festplatten) mit Ethernet Ports ausgestattet, mit dem alle Sensoren verbunden sind. Die Stromversorgung der Kameras, der Beleuchtungseinheiten und des PCs wird über einen Generator gewährleistet.

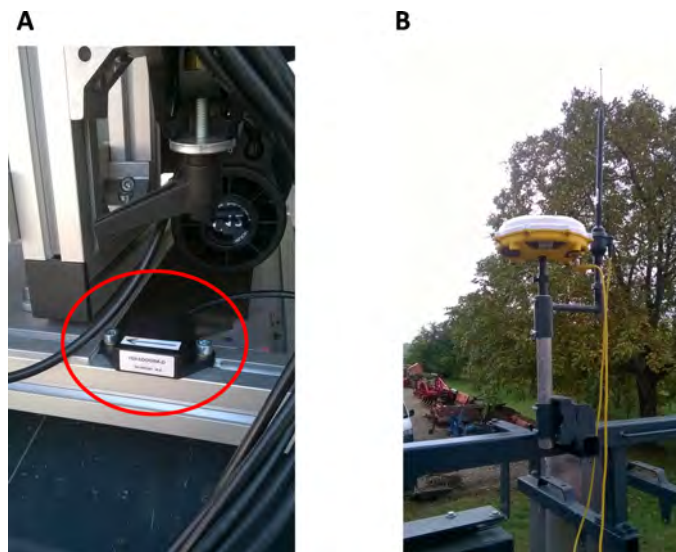


Abbildung 5.1: **A:** Das Inklinometer (rot markiert) ist parallel zur Längsachse des Fahrzeuges montiert und bestimmt dessen Roll- und Pitch-Rotation. **B:** Die Antenne des RTK-GPS ist auf dem oberen Fahrzeugrahmen montiert. Bildautor: Markus Wieland

5.1.1.2 Vertikales Multi-Kamerasystem

Das Multi-Kamerasystem besteht aus drei vertikal und parallel zueinander angeordneten Industriekameras (DALSA Genie NanoC2590 (DALSA Inc., Waterloo, ON, Canada, $u = 2592 \times v = 2048\text{pel}$; horizontale aktive Sensorlänge $12,4\text{mm}$, vertikale aktive Sensorlänge $9,8\text{mm}$; Objektiv: KOWA LM12HC 12mm Brennweite). Abbildung 5.2 zeigt die Konfiguration des MKS. Alle Kameras sind in Schutzgehäusen (silbern mit grünen Endteilen) montiert und werden beginnend mit der obersten Kamera mit Nummer 1 – 3 bezeichnet.

Sechs Halogenstrahler (Hedler C12, 300 Watt , Hedler Systemlicht, Runkel/Lahn, Deutschland) sind um das MKS zur gleichmäßigen Beleuchtung der Szenerie platziert. Sie emittieren ein Licht mit warmer Farbtemperatur (3200 Kelvin). Die Bilder besitzen daher einen Gelbstich. Es ist jedoch zu erwarten, dass für die unter

anderem auf Farbmerkmalen basierte Klassifikation unerheblich ist, da die relativen farblichen Unterschiede der Pflanzenteile gewahrt bleiben.

Alle Kameras sind mit dem PC über ein GigE Interface verbunden. Jede Kamera ist dabei einem eigenen Ethernet Port zugewiesen und mit Schnittstellen für Automatisierungs- und Triggerprozesse für die zeitliche Steuerung mehrerer Kameras ausgestattet.

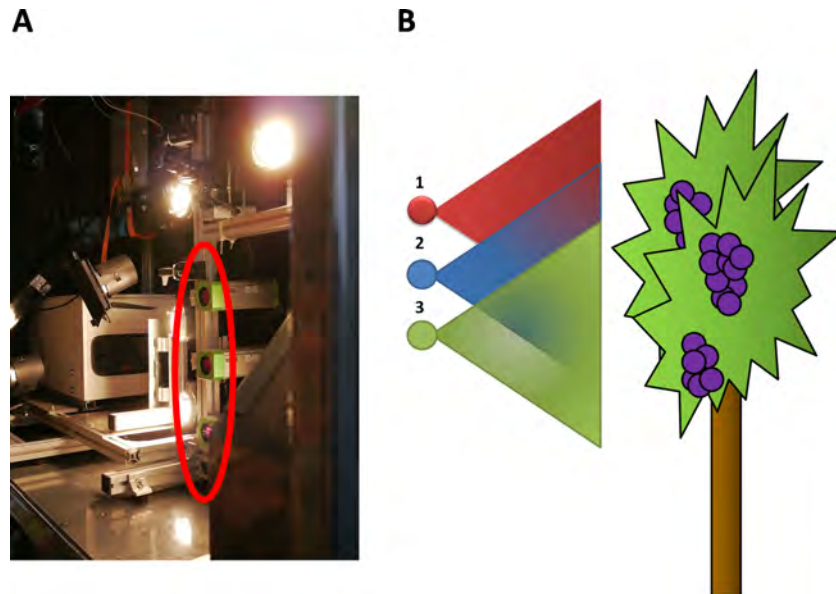


Abbildung 5.2: **A:** Das MKS (rot markiert) mit rundherum angebrachten Beleuchtungseinheiten. **B:** Die vertikale Bildüberlappung der Kameras 1-3 (rot, blau, grün) wird durch die Anordnung der Kameras erreicht.

Die für MVS nötige Bildüberlappung wird vertikal durch die parallele Anordnung der Kameras gewährleistet. Durch die parallele Anordnung entlang der Vertikalen soll eine maximale Bildabdeckung der Weinreben bei maximaler Bildüberlappung erreicht werden. Horizontal wird die Überlappung durch die Bewegung des Phenoliners parallel zu den Anbaureihen erreicht. Kurze Belichtungszeiten von $5ms$ verhindern Bewegungsunschärfe in den Bildern bei Bildaufnahmen aus der Bewegung. Alle Kameras sind vom gleichem Modell, mit identischem Objektiv und gleicher Fokuseinstellung. Sie besitzen daher annähernd identische innere Kameraparameter, so dass die Bilder ebenfalls annähernd identische Eigenschaften besitzen. Dies ist für die spätere Rekonstruktion durch Pix4DMapper wichtig, da die Software von lediglich einer Kamera mit konstanten inneren Kameraparametern ausgeht. Abbildung 5.3 skizziert die Bildaufnahme mit dem Phenoliner.

Bei der Berechnung der Bildüberlappung wird nach den Formeln aus Kapitel 4.1.2 der Voruntersuchungen vorgegangen. Das MKS des Phenoliners fährt in einem Abstand H von ca. $750mm$ an den Weinreihen vorbei. Die horizontale Länge S der horizontalen Bildachse beträgt $12,4mm$, die Kamerakonstante F beträgt $12mm$. Die projizierte horizontale Strecke S_O , die der Bildsensor entlang seiner horizontalen Bildachse im Objektraum erfasst, beträgt nach Formel 4.1 ca. $780mm$.

Die GSD des Phenoliners beträgt demnach mit der horizontalen Pixelauflösung u von $2592pel$:

$$GSD = \frac{780mm}{2592} \approx 0,3mm. \quad (5.1)$$

Die horizontale Bildüberlappung wird durch die Bewegung des Phenoliners erreicht. Im Unterschied zum PHENObot wird die Bildaufnahme nicht manuell ausgelöst, sondern geschieht automatisch mit einer Fre-

quenz von $5Hz$ (s. Kapitel 5.1.1.3). Bei einer Geschwindigkeit $v = 0,6km/h \approx 167mm/s$ beträgt der horizontale Abstand $base_h$ zweier Bildern zwischen zwei Aufnahmezeitpunkten demnach

$$\begin{aligned} base_h &= \frac{1}{(5 \cdot v)} \\ &= 33mm. \end{aligned} \quad (5.2)$$

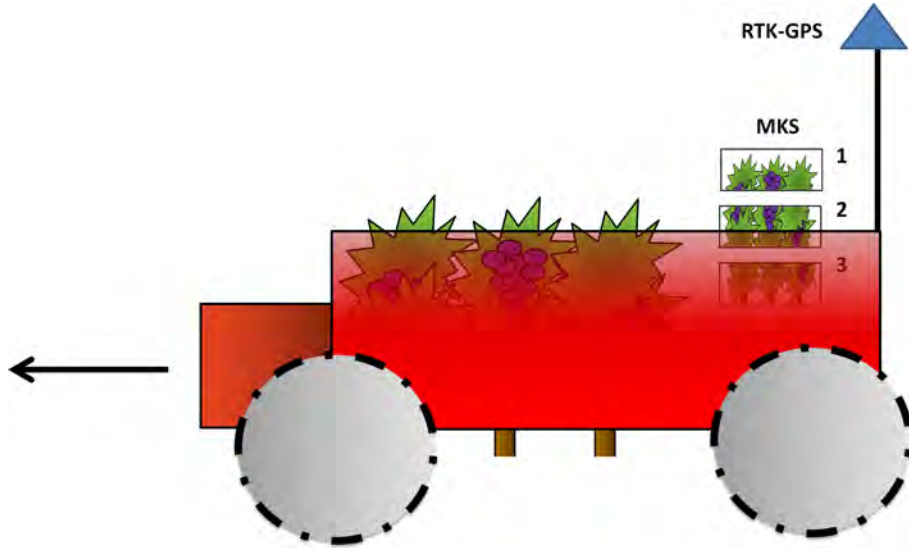


Abbildung 5.3: Skizze der Datenaufnahme mit dem Phenoliner. Der Tunnel wird über den Anbaureihen platziert, hier durch die Transparenz des Phenoliners angedeutet.

Die Bildüberlappung o_h beträgt nach Formel 4.2 bei der horizontalen S_O von $780mm$ somit

$$\begin{aligned} o_h &= 1 - \frac{33}{780} \\ &= 96\%. \end{aligned} \quad (5.3)$$

Mit der projizierten vertikalen Strecke S_O kann der maximale vertikale Abstand $base_v$ der einzelnen Kameras zueinander berechnet werden, um eine bestimmte vertikale Bildüberlappung o_v zu gewährleisten. Nach Formel 4.1 ergibt sich für die vertikale S_O einer Kamera des Phenoliners:

$$S_O = 612mm$$

Um die Rekonstruktion zu verbessern, wird vertikal eine höhere Überlappung o_v als beim PHENObot von mind. 75% angestrebt. Der notwendige vertikale Abstand $base_v$ ergibt sich für $o_v = 75\%$ nach Formel 4.3:

$$\begin{aligned} base_v &= (1 - 0,75) \cdot 612mm \\ &= 153mm. \end{aligned} \quad (5.4)$$

Die drei Kameras wurden folglich in einem vertikalen Abstand von $base_v = 153mm$ zueinander montiert. Vertikal wird nach Formel 4.4 damit insgesamt eine Strecke von $S_v = 918mm$ abgedeckt.

5.1.1.3 IggGeotagger 2.0

Erfasste Sensordaten

Die Software IggGeotagger 2.0¹ [Kicherer et al., 2017] kontrolliert die Aufnahme und Speicherung der Bilder,

¹Autor des IggGeotagger 2.0 ist Markus Wieland, IGG Bonn

der GPS-Position und der Rotationsdaten des Inklinometers und führt die Georeferenzierung der Bilder durch. Sie ist die Nachfolgerversion der Software IggGeotagger, die bereits im PHENObot genutzt wurde. Die Aufnahme der Bilder, sowie der GPS- und Inklinometer Daten erfolgt nach Aktivierung des IggGeotagger vollständig automatisch. Alle Daten der genannten Sensoren werden auf den SSD-Festplatten (Solid State Disc) des PCs abgelegt. Die Bildaufnahme erfolgt für alle Kameras zeitsynchron mit einer nutzerdefinierten Frequenz von maximal $5Hz$, welche durch die Geschwindigkeit der Datenspeicherung limitiert ist. Die GPS-Positionen werden als geographische Koordinaten mit Längen- und Breitengraden, sowie der ellipsoidischer Höhe in einem NMEA-String (National Marine Electronics Association) mit einer Frequenz von $20Hz$ gemessen. Roll- und Pitch-Rotation des Fahrzeugs werden vom Inklinometer mit einer Frequenz von $20Hz$ aufgezeichnet.

Alle Daten der einzelnen Sensoren erhalten durch den IggGeotagger 2.0 einen Zeitstempel der Systemzeit des PCs zum Aufnahmezeitpunkt. Über diesen Zeitstempel werden korrespondierende Messungen aller Sensordaten zeitlich zueinander zugeordnet.

Georeferenzierung der Bilder im Postprocessing

Im Postprocessing werden alle Sensordaten vom IggGeotagger 2.0 zur Georeferenzierung der Bilder genutzt. Basierend auf den zeitidentischen Sensordaten werden im Postprocessing die Position jedes Bildes sowohl in geographischen Koordinaten mit ellipsoidischer Höhe, sowie als verebnete *UTM-Koordinaten* (*Universal Transverse Mercator*) mit Rechtswert X und Hochwert Y bestimmt. Dabei müssen die relativen Abstände zwischen GPS-Antenne und den Kameras des MKS, sowie die Neigung des Fahrzeugs entlang seiner Achsen berücksichtigt werden. Der relative Abstand zwischen GPS-Antenne und Kamera wird als Hebelarm h bezeichnet. Um ihn zu bestimmen, wurde der Phenoliner von außen und aus dem Tunnel heraus mit einem terrestrischen Laserscanner (Leica P20, Leica Microsystems GmbH, Wetzlar, Deutschland) gescannt, wobei die einzelnen Scans über Zielzeichen zueinander registriert wurden (Abbildung 5.4). Anhand der Punktwolke des Phenoliners wurde der Hebelarm für die unterste Kamera 3 bestimmt.

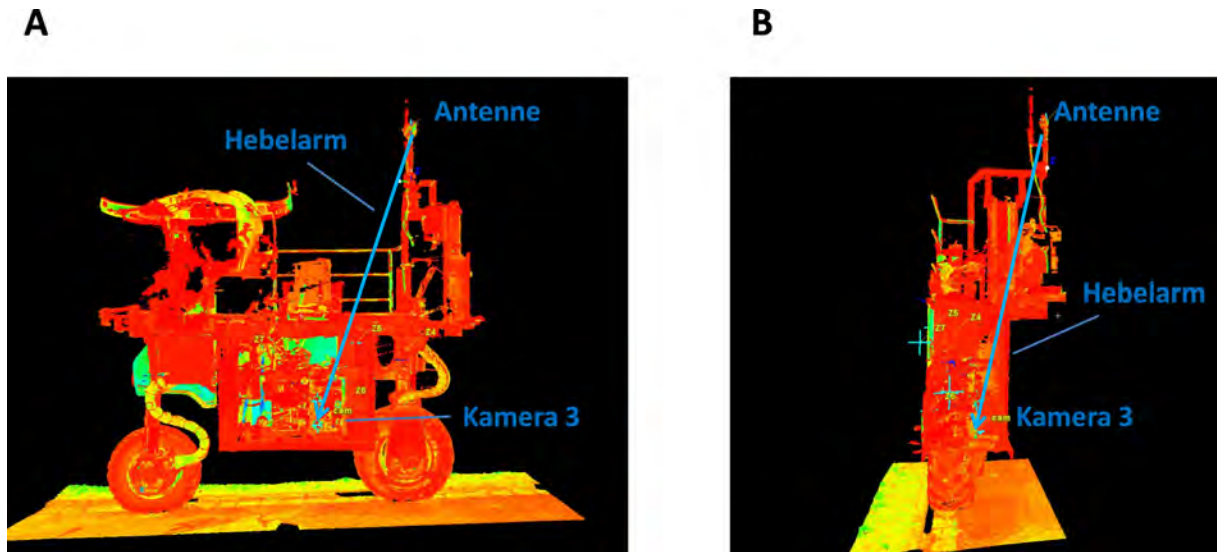


Abbildung 5.4: **A:** Seitliche Ansicht des 3D-Scans, gefärbt nach Reflektionsintensitäten des Messsignals. **B:** Ansicht des 3D-Scans vom Heck des Phenoliners. Hebelarm, Position der Antenne und der Kamera 3 sind eingezeichnet. Bilder der Scans: Markus Wieland. Markierungen durch den Autor.

Die Positionen der GPS-Antenne $[GPS_X, GPS_Y, GPS_Z]$ und der Kamera 3 $[MKS_X, MKS_Y, MKS_Z]$ sind im lokalen Koordinatensystems des Laserscanners angegeben. Für den Hebelarm der Kamera 3 h_3 gilt

$$\mathbf{h}_3 = \begin{bmatrix} \Delta(GPS_X, MKS_X) \\ \Delta(GPS_Y, MKS_Y) \\ \Delta(GPS_Z, MKS_Z) \end{bmatrix}. \quad (5.5)$$

Der Kurswinkel des Phenoliners um seine Hochachse (Yaw) im Bezug zum UTM-System wird aus der Fahrzeugposition $\mathbf{p}_t = [X, Y, Z]_t$ zu unterschiedlichen Zeitpunkten t_1 und t_2 berechnet. Die Fahrzeugposition ist dabei synonym mit der Position der GPS-Antenne.

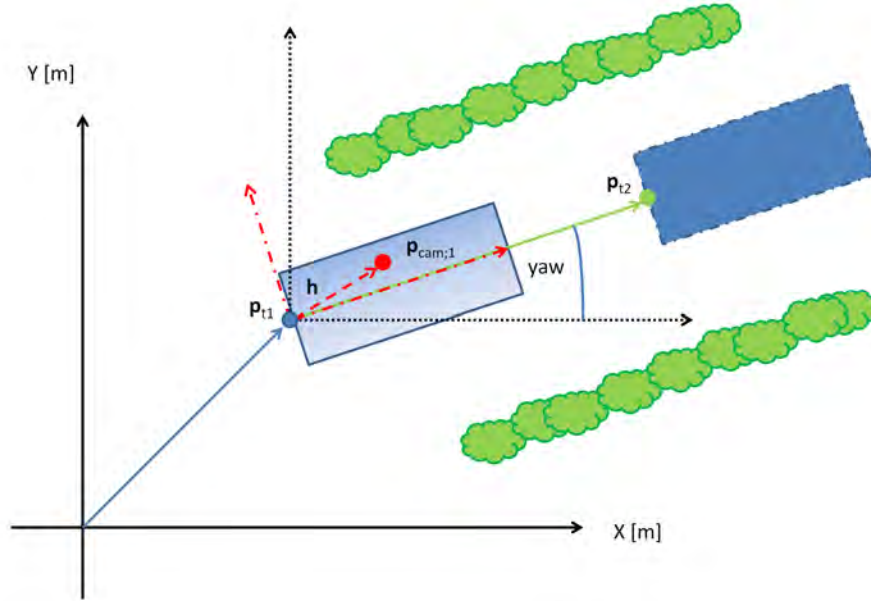


Abbildung 5.5: Der Kurswinkel des Phenoliners wird aus der Fahrzeugposition $\mathbf{p}_t$ zu zwei unterschiedlichen Zeitpunkten t_1 und t_2 bestimmt. Rote Pfeile (punkt-gestrichelt) stellen Längs- und Querachse des Phenoliners dar, der grüne Pfeil den Vektor zwischen beiden Fahrzeugpositionen $\mathbf{p}_{t1}$ (blauer Kreis) und $\mathbf{p}_{t2}$ (grüner Kreis) zu Zeitpunkten t_1 und t_2 , der blaue Pfeil den Ortsvektor $\mathbf{p}_{t1}$. Der rot-gestrichelte Pfeil h entspricht dem Hebelarm einer Kamera. Mit $\mathbf{p}_{cam;1}$ sei die Position der Kamera im UTM-System bezeichnet.

Der Kurswinkel ergibt sich über

$$yaw = atan \left(\frac{\Delta(Y_{t2}, Y_{t1})}{\Delta(X_{t2}, X_{t1})} \right). \quad (5.6)$$

Die Berechnung des Kurswinkels aus unterschiedlichen Positionen zu zwei Zeitpunkten hat zur Folge, dass die Bewegung des Phenoliners durch die Anbaureihen während der Datenaufnahme auf eine lineare Vorwärtsbewegung entlang seiner Längsachse modelliert ist. Da die Anbaureihen geradlinige Pflanzungen sind, existiert kein Widerspruch zwischen Modellierung und tatsächlicher Bewegung.

Der horizontale Abstand zweier Bilder beträgt wie beschrieben rund $3cm$ und liegt damit ungefähr innerhalb der Unsicherheit der RTK-GPS Positionsbestimmung ($\sigma = 2cm$). Damit letztere keinen zu großen Einfluss auf die Winkelberechnung ausübt, muss der horizontale Abstand der Bilder zu t_1 und t_2 ausreichend groß genug gewählt werden. Dieser Umstand wird im IggGeotagger 2.0 entsprechend berücksichtigt. Mit dem Hebelarm $\mathbf{h}_3$ und den Yaw-, Pitch- und Roll-Rotationswinkeln werden die Koordinaten $\mathbf{p}_{cam;3}$ der Kamera 3 im UTM-System zum Zeitpunkt t_1 berechnet. Die Rotation des Hebelarms um die einzelnen Achsen erfolgen in der Reihenfolge Roll, Pitch, Yaw. Sie werden einzeln durch Rotationsmatrizen $\mathbf{R}_{roll}$, $\mathbf{R}_{pitch}$, $\mathbf{R}_{yaw}$ angegeben. Diese lauten:

$$\mathbf{R}_{roll} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(roll) & -\sin(roll) \\ 0 & \sin(roll) & \cos(roll) \end{bmatrix} \quad (5.7)$$

$$\mathbf{R}_{pitch} = \begin{bmatrix} \cos(pitch) & 0 & \sin(pitch) \\ 0 & 1 & 0 \\ -\sin(pitch) & 0 & \cos(pitch) \end{bmatrix} \quad (5.8)$$

$$\mathbf{R}_{yaw} = \begin{bmatrix} \cos(yaw) & -\sin(yaw) & 0 \\ \sin(yaw) & \cos(yaw) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.9)$$

Die Rotationen erfolgen dabei gegen den Uhrzeigersinn positiv. Die vollständige Rotationsmatrix ergibt sich folgendermaßen:

$$\mathbf{R} = \mathbf{R}_{yaw} \cdot \mathbf{R}_{pitch} \cdot \mathbf{R}_{roll}. \quad (5.10)$$

Die Position der Kamera 3 im UTM-System wird folgendermaßen berechnet:

$$\mathbf{p}_{cam;3} = \mathbf{p}_{t1} + \mathbf{R} \cdot \mathbf{h}_3. \quad (5.11)$$

Aufgrund der parallelen, vertikalen Anordnung der Kameras wird für die Kameras 1 und 2 von identischen Positionen in X- und Y-Koordinaten wie für Kamera 3 ausgegangen und ihnen diese Koordinaten zugewiesen. Die Z-Koordinaten der Kameras 1 und 2 ergeben sich aus den vertikalen Abständen $base_v$ zwischen den drei Kameras bezogen auf die Z-Koordinate von Kamera 3. Die UTM-Position aller Kameras des MKS werden zuletzt zurück in geographische Koordinaten transformiert und im Exif-Header des entsprechenden Bildes gespeichert.

5.1.2 Bildaufnahme mit dem Phenoliner

5.1.2.1 Betrachtete Weinreben

Mit dem Phenoliner wird jeweils eine Riesling Anbaureihe nach dem TS- und nach dem MSS-Erziehungssystem im BBCH89 Reifestadium in der Saison 2016 erfasst. Für die Schätzung der Ertragsparameter und des Ertrages ist die Erhebung von Referenzwerten notwendig (näheres dazu im Kapitel Quantifikation). Die Referenzwerte werden im späteren Verlauf in jeweils 5x1-Meter Abschnitten der TS- und MSS-Anbaureihe erhoben. Alle 1-Meter Abschnitte wurden zu diesem Zweck jeweils mit Flutterband abgegrenzt und in jedem Abschnitt jeweils fünf Trauben, im Weiteren als *Referenztrauben* bezeichnet, anhand von durchlaufend nummerierten Zetteln markiert. Insgesamt ergeben sich somit 25 Referenztrauben pro Anbaureihe. Abschnitte und Referenztrauben sind auf diese Weise auf den Bildern und damit in den Punktwolken eindeutig identifizierbar. Das Blattwerk der fünf Abschnitte der TS-Anbaureihe ist in Traubennähe leicht ausgedünnt worden, was einer gängigen Methodik für TS-Erziehungssysteme entspricht [Reynolds et al., 1986]. Das Blattwerk der MSS-Anbaureihe wurde nicht verändert.

5.1.2.2 Vorgehensweise, Arbeitsaufwand und Durchsatz

Abbildung 5.6 zeigen den Phenoliner bei der Bildaufnahme im Feldeinsatz. Der Tunnel des Phenoliners wird zur Datenerfassung über den Anbaureihen positioniert, wobei sich seine linken und rechten Reifen im jeweils linken bzw. rechten Korridor zwischen zwei Anbaureihen befinden. Das MKS ist dabei parallel zu den Anbaureihen ausgerichtet. Die Anbaureihen von 25m Länge wurden bei einer Geschwindigkeit von 0,6km/h in einer Zeit von etwa zweieinhalb Minuten einmalig durchfahren und vom MKS bildlich erfasst. Alle drei Kameras werden über die Schnittstellen in nur einer einzigen Zeiteinheit [Weinmann et al., 2011] simultan ausgelöst. Die simultane Auslösung aller Kameras ist besonders für den Fall der dynamischen Weinreben wichtig. Leichte Bewegungen der Weinrebe in wenigen Bildern können daher durch die Verarbeitung im Bildverbund und den identischen Zeitbezug kompensiert werden, falls das Objekt im Großteil der Bilder dennoch unbewegt verbleibt. Der Arbeitsaufwand pro Anbaureihe ist damit um ein Vielfaches geringer als mit dem PHENObot (15 Minuten) und erfüllt die Anforderungen an den Hochdurchsatz.



Abbildung 5.6: Der Phenoliner im Einsatz. Die Weinreben befinden sich im Tunnel, das Multi-Kamerasystem wird seitlich durch die Anbaureihen bewegt. Bildautor: Vine Xpert Digitaler-Weinbau GmbH.

Die Anbaureihen haben eine Länge von $25m$. Ziel ist es, an einem Tag 120 Anbaureihen erfassen zu können. Bei einer durchschnittlichen Länge einer Anbaureihe von $25m$ entspricht dies einer Strecke von $3km$. Zusätzlich wird je Anbaureihe ein Zeitzusatz von 3 Minuten $= 1/20[h]$ für Wendemanöver und das Wechseln des Phenoliners in die nächste Anbaureihe addiert. Der Arbeitsaufwand in Stunden $[h]$ zur Erfassung der 120 Anbaureihen ergibt sich folgendermaßen:

$$\begin{aligned} \text{Arbeitsaufwand}[h] &= \frac{3km}{0,6km/h} + \frac{1}{20}[h] \cdot 120 \\ &= 5[h] + 6[h] \\ &= 11[h] \end{aligned} \tag{5.12}$$

Für die Erfassung der 120 Anbaureihen werden nach diesen Überlegungen 11 Stunden veranschlagt. Das Ziel, 120 Anbaureihen an einem Tag mit dem Phenoliner erfassen zu können, ist also erreichbar. Je Anbaureihe werden etwa 2500 Bilder im JPG-Format mit 1,5MB aufgenommen. Dies stellt einen Speicherbedarf von ca. 10GB pro Reihe dar.

5.1.2.3 Bildabdeckung, -anzahl und -qualität

Im TS-Erziehungssystem wachsen die Trauben dicht über dem Boden und in primär einem Höhenniveau. Die Höhe des Phenoliners wurde für die TS-Anbaureihe auf die durchschnittliche Höhe der Trauben angepasst. Es kann daher davon ausgegangen werden, dass die vertikale Bildabdeckung des Phenoliners bei TS-Anbaureihen vertikal alle sichtbaren Trauben erfasst. Im MSS-Erziehungssystem wachsen Trauben über die gesamte Tiefe und Höhe der Weinrebe verteilt. Die vertikale Ausdehnung des Blattwerks beträgt etwa $110 - 130cm$. Mit der vertikalen Bildabdeckung des MKS von $S_v = 91,8cm$ werden die Trauben daher zu einem großen Teil, jedoch nicht vollständig erfasst.

Pro Reihe werden etwa 2500 Bilder mit jeweils $15Mb$ (ca. $37GB$ pro Anbaureihe) aufgenommen. Abbildungen 5.7 und 5.8 zeigen jeweils einen 1-Meter Abschnitt der TS- bzw. MSS-Anbaureihe mit markierten Referenztrauben.



Abbildung 5.7: Betrachtete Riesling Reihe nach der TS-Erziehung. Referenztrauben, die später zur Erhebung von Referenzwerten benutzt werden, wurden durch nummerierte Zettel für ihre Wiedererkennung markiert. Der weiße Hintergrund füllt ungefähr ein Drittel des Bildes, wird jedoch aufgrund der Farbtemperatur der Halogenstrahler nicht weiß, sondern gelblich abgebildet. Das Blattwerk wurde für die fünf Abschnitte manuell ausgedünnt. (Anmerkung: Einzelne Beeren sind durch runde Aufkleber markiert. Diese Markierungen sind jedoch für die hier geschilderten Untersuchungen nicht relevant.)

In den Bildern ist der durch die Farbtemperatur der Halogenstrahler verursachte Gelbstich zu beobachten. Der weiße Hintergrund reflektiert stark und wird in einem gelben Farbton dargestellt. Zusätzlich lässt sich Schattenwurf auf dem Hintergrund erkennen. Unterschiedliche Tiefen des Blattwerks werden unterschiedlich intensiv erleuchtet. Eine durchgehend gleichmäßige Beleuchtung ist daher nicht möglich.

Diesbezüglich hat vor allem die Morphologie der Weinreben in beiden Erziehungssystemen Auswirkungen auf die Intensität der Beleuchtung und die Auswertbarkeit der Bilder. Auf der Mitte der Beeren der TS-Anbaureihe können bei näher an den Beleuchtungseinheiten liegende Trauben gelegentlich überbelichtete Stellen beobachten werden. Diese Stellen besitzen aufgrund der Überbelichtung homogene RGB-Werte und unterscheiden sich farblich von den nicht überbelichteten Stellen am Rand der Beeren, was die Zuordnung homologer Punkte für die Rekonstruktion erschwert.



Abbildung 5.8: Betrachtete Riesling Reihe nach der MSS-Erziehung. Auch hier wurden Referenztrauben für ihre spätere Wiedererkennung markiert. Durch die hohe Dichte des Blattwerks ist der Hintergrund nur an wenigen Stellen sichtbar. Primär Trauben im Vordergrund häufig überbelichtet.

Im Falle der MSS-Anbaureihen ist das Phänomen der Überbelichtung noch häufiger zu beobachten. Größere Teile von Trauben sind überbelichtet. Dies hat zwei Gründe: 1. Trauben wachsen in der MSS-Erziehung über die gesamte Tiefe und Höhe der Weinrebe verteilt. Mehr Trauben befinden sich daher im Vergleich zur TS-Erziehung näher an den Beleuchtungseinheiten. 2. Aufgrund der großen Tiefe des Blattwerks muss der Tunnel zwangsweise mittig zu den Anbaureihen positioniert werden, um das vollständige Volumen des Blattwerks aufnehmen zu können. Die Distanz zwischen Beleuchtungseinheiten und Weinreben wird dadurch geringer. Im Gegensatz dazu können bei der TS-Anbaureihe aufgrund der geringeren Tiefe der Weinreben das MKS und die Beleuchtungseinheiten eine größere Distanz zur betrachteten Anbaureihe einnehmen. Der Unterschied in der Distanz zwischen beiden Erziehungssystemen ist in den Abbildungen bspw. an der Dicke der Äste oder der Beerengröße erkennbar.



Abbildung 5.9: Ein herausragender Ast (rote Pfeile) verdeckt den Blick auf die Weintrauben in drei sequentiell aufeinanderfolgenden Bildern und wird durch Kontakt mit dem Phenoliner bewegt. Im mittleren Bild verlässt das zugehörige Blatt den Bildbereich ganz.

Das größere Volumen des Blattwerks der MSS-Anbaureihe führt zu einem weiteren Phänomen, dem Kontakt zwischen herausragenden Ästen und dem Phenoliner. Durch die Berührung kann es zu Bewegungen der Weinreben während der Bildaufnahme kommen. Herausragende Äste ragen stellenweise direkt vor die Kameralinse und verdecken den Blick der Kamera auf die Weinreben, wie Abbildung 5.9 zeigt. Durch die Dynamik der homologen Punkte an diesem Ast wird die Bestimmung der relativen Kamerapositionen und die Triangulation verfälscht (s. folgendes Kapitel).

5.1.3 3D-Rekonstruktion der Anbaureihen

5.1.3.1 Georeferenzierung der Bilder

Abbildung 5.10 zeigt einen Ausschnitt der MSS-Anbaureihe mit initialen und während durch Pix4DMapper geschätzten Kamerapositionen. Die vertikale Anordnung der Kameras des MKS ist anhand der Anordnung der grünen Kugeln gut erkennbar. Die Abweichung zwischen geschätzter Kameraposition und initialer GPS-Position betrug für die TS- bzw. MSS-Anbaureihe im Durchschnitt 2,5cm für den Rechtswert, 2,4 bzw. 2,7cm für den Hochwert und 0,8 bzw. 1,2cm für die Höhe. Im Vergleich zum PHENObot sind die Residuen durch die Berücksichtigung des Hebelarms zwischen GPS-Antenne und MKS, sowie der Fahrzeugneigung deutlich geringer. Die Werte liegen im Rahmen der spezifizierten Unsicherheit der RTK-GPS Positionsbestimmung von 2cm und korrespondieren mit den Ergebnissen der Georeferenzierung aus Kicherer et al. [2017], wo die Bildpositionen mit Soll-Koordinaten der abgebildeten Weinreben verglichen wurden.

5.1.3.2 Zeitaufwand

Die Rekonstruktion einer Anbaureihe aus etwa 2500 Bildern dauert an einem PC mit 12Gb Arbeitsspeicher 20–22 Stunden. Die enorme zu verarbeitende und zu organisierende Datenmenge erzeugt neue Flaschenhälse in der Phänotypisierung [Araus and Cairns, 2014; Minervini et al., 2015]. Die Verwendung von entsprechend großen und schnellen Speichermedien, sowie leistungsstarker Recheneinheiten ist daher unerlässlich für eine zeitnahe Verarbeitung der Daten.

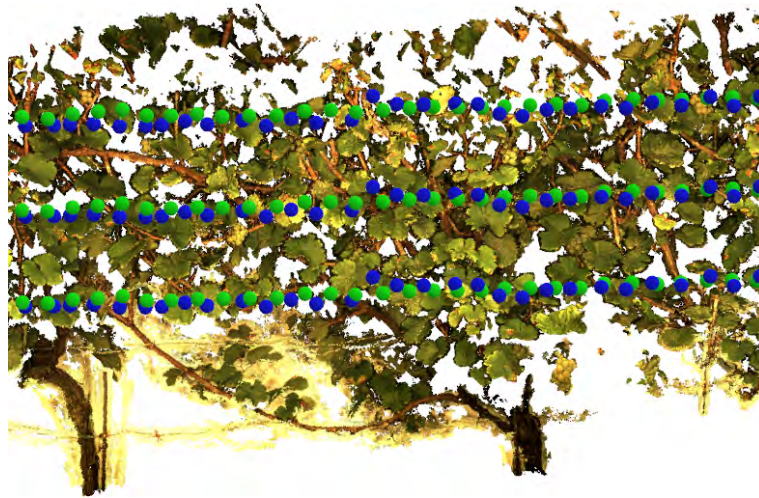


Abbildung 5.10: Ausschnitt der MSS-Anbaureihe. Grüne Kugeln indizieren geschätzte Kamerapositionen $\mathbf{p}_{cam}$ nach der Bündelblockausgleichung, blaue Kugeln die initialen GPS-Positionen der Kameras aus der Georeferenzierung. Die Residuen zwischen initialer und geschätzter Kameraposition sind klein.

5.1.3.3 Genauigkeit der Rekonstruktion

Der durchschnittliche Rückprojektionsfehler d der gesamten rekonstruierten 25m langen Anbaureihen beträgt 0,2pel für die TS-Anbaureihe und 0,3pel für die MSS-Anbaureihe (PHENObot 0,1 – 0,2pel). Die Bedingung $d \ll 1pel$ ist damit für beide Anbaureihen erfüllt. Im Vergleich zu den Untersuchungen mit dem PHENObot ist d leicht gestiegen. Aus den Bildern des Phenoliner werden vermehrt Punkte des Hintergrundes rekonstruiert (s. Abbildung 5.11), an denen durch Schattenwurf und Bewegung des Tuchs heterogene Bildinhalte entstanden und die daher rekonstruiert werden konnten. Da diese Bildinhalte lediglich durch Schattenwurf entstehen, stellen sie kein tatsächliches, formfestes geometrisches Objekt dar und tragen zu einer leichten Erhöhung des Rückprojektionsfehlers bei.

Die Punkt-zu-Punkt-Distanz der Punktwolke beträgt 0,3 – 0,5mm und entspricht damit ungefähr der GSD des MKS. Die innere Genauigkeit und die geometrische Auflösung der Rekonstruktion ist damit wie auch beim PHENObot hoch genug, die runde Oberfläche der kleinmaßstäbigen Beeren von wenigen Millimetern Durchmesser korrekt darzustellen.

Im Kapitel 4.1.3 zur Rekonstruktion aus Bildern des PHENObots wurden die Punktwolken beider Erziehungssysteme anhand der mittleren Varianz des Rückprojektionsfehlers d und der mittleren Anzahl von Bildern bewertet, in denen ein Objektpunkt beobachtet wurde. Um einen Vergleich zum PHENObot zu ziehen, werden diese Größen auch für den Phenoliner bestimmt und mit den vorherig bestimmten Werten des PHENObots verglichen. Für die Berechnung wird identisch zu Kapitel 4.1.3 vorgegangen. Insgesamt werden jeweils 11 3D-Punkte auf vier Trauben im Pix4DMapper Editor betrachtet und die beiden Größen abgelesen. Tabelle 5.1 listet die mittlere Anzahl der Beobachtungen eines Punktes, sowie die mittlere Varianz des Rückprojektionsfehlers der 44 einzelnen 3D-Punkten auf.

	$\overline{Bilder}$	$\overline{\sigma_d^2}$ [pel]
TS-Anbaureihe	59	1,1
MSS-Anbaureihe	26	2,4

Tabelle 5.1: Mittlere Anzahl der Beobachtungen und mittlere Varianz des Rückprojektionsfehlers d aus 44 3D-Punkten beim Phenoliner.

Identisch zu den Voruntersuchungen werden Trauben der TS-Anbaureihe häufiger beobachtet, als Trauben der MSS-Anbaureihe (vgl. Tabelle 4.1). Die mittlere Varianz des Rückprojektionsfehlers ist für die TS-Anbaureihe geringer als für die MSS-Anbaureihe. Dies kann einerseits mit dem höheren Verdeckungsgrad der Trauben in der MSS-Anbaureihe erklärt werden, andererseits durch die erwähnte Überbelichtung resultieren, die homogene Farbwerte erzeugt und dadurch die Zuordnung homologer Bildpunkt verhindert.

Es lässt sich im Vergleich zur Bildaufnahme mit dem PHENObot zudem feststellen, dass sich durch die Anordnung des MKS, sowie die häufigere Bildaufnahme auch die Anzahl der Beobachtungen eines Punktes von 21 Bildbeobachtungen auf 59 Bildbeobachtungen bei der TS-Anbaureihe und von 18 zu 26 Bildbeobachtungen bei der MSS-Anbaureihe verbessert hat. Die relative Steigerung der Beobachtungsanzahl weist auf eine allgemein erhöhte Wahrscheinlichkeit hin, dass ein Punkt zur Rekonstruktion häufig genug beobachtet werden kann.

Zugleich ist die mittlere Varianz des Rückprojektionsfehlers einzelner Traubenpunkte im Vergleich zum PHENObot gestiegen (TS 0,2pel, MSS 0,7pel). Dies hat mehrere Ursachen.

Durch Äste induzierte Bewegungen der Weinreben, die in Folge nicht korrekt rekonstruiert wurden, können zu einem erhöhten Rückprojektionsfehler beitragen. Eine weitere Ursache ist die Überbelichtung, von der vor allem die MSS-Anbaureihe betroffen ist und welche die korrekte Zuordnung homologer Punkte verhindert. Die Aufnahmegeometrie der Bilder besitzt ebenfalls Einfluss auf die Rekonstruktion. Da die Bilder im Phenoliner mit einem horizontalen Abstand $base_h$ von ca. 3cm (PHENObot: $base_h = 10 - 15cm$) aufgenommen werden, verringert sich der Schnittwinkel α , unter dem sich die Projektionsstrahlen der Bilder schneiden. Die dadurch entstehenden schleifenden Schnitte können die Unsicherheit der Triangulation erhöhen.

5.1.4 Vorverarbeitung der Punktwolken

Die Vorverarbeitungsschritte sind identisch wie im vorherigen Kapitel der Voruntersuchungen. Auf detaillierte Erläuterungen der Funktionsweise der einzelnen Schritte wird an dieser Stelle verzichtet. Die Schritte umfassen die Entfernung von Punkten des Hintergrundes, das Entfernen von Ausreißern, das Glätten der Punktwolke sowie die Verringerung der Punkt-zu-Punkt Distanz auf 1mm.

Die weiße Farbe der linken Tunnelwand wurde ausgewählt, da dieser Farbbereich keiner möglichen Färbung einer Traubensorte entspricht und weiß rekonstruierte 3D-Punkte daher mittels Schwellwertdiskriminierung entfernt werden können, ohne dabei Traubenpunkte zu löschen. Der eigentlich weiße Hintergrund ist durch die Belichtung jedoch hellgelb gefärbt. Punkte dieser Färbung sind zu entfernende Punkte des Hintergrunds.

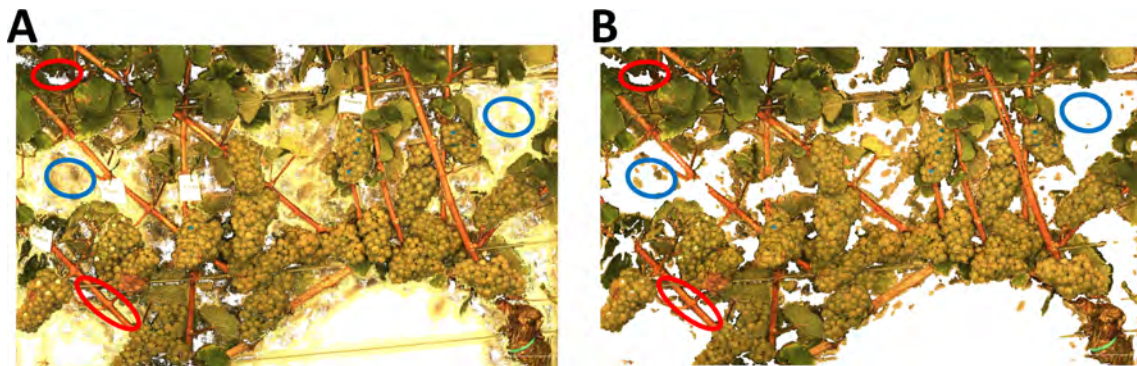


Abbildung 5.11: **A:** Ausschnitt aus einer Punktwolke. Punkte an Objekträndern entstehen durch Mischpixel aus Vorder- und Hintergrund (Beispiele rot markiert). Ungeordnete, dünne Punktreionen des Hintergrunds (Beispiele blau markiert). **B:** Der Ausschnitt nach der Entfernung des Hintergrundes. Die falsch triangulierten Punkte wurden größtenteils entfernt. Kleinere Regionen dieser Punkte verbleiben jedoch weiterhin in der Punktwolke. Diese Regionen entsprechen zum Großteil Stellen auf dem Hintergrund, wo die hellgelbe Färbung durch Schattenwurf verdunkelt ist.

Abbildung 5.11 zeigt einen unbearbeitenden Punktwolkenausschnitt, der falsch triangulierte Punkte des Hintergrundes enthält sowie denselben Ausschnitt nach der Vorverarbeitung. Gut erkennbar sind die hellgelben,

sowie die sehr dunklen Punkte vor allem am Blattwerk, welche die Objektränder verdicken und durch die Triangulation von Mischpixeln aus Vorder- und Hintergrund entstehen. In der Nähe heterogen gefärbter Objekte entstehen des Weiteren ungeordnete, dünne Punktregionen des Hintergrunds.

Der Hintergrund wird nach der in Kapitel 4.1.4 beschriebenen Methode über Konvertierung in den HSV-Raum entfernt. Im Falle des eigentlich weißen Hintergrunds muss die im Bild dargestellte Farbe anstelle der tatsächlich weißen Farbe ermittelt werden. Aus mehreren Punkten des Hintergrunds werden dazu empirisch HSV-Durchschnittswerte ermittelt. Die aufgrund seiner stark reflektierenden weißen Farbe hohe Helligkeit des Hintergrunds ist dabei ein herausragendes Merkmal. Als Punkte des Hintergrunds werden demnach Punkte entfernt, die einen Helligkeitswert von $V = 0,98$ überschreiten, sowie einen Saturation-Wert von $S = 0,63$ unterschreiten. Die gewählten Schwellwerte löschen effektiv die Punkte des weißen, stark überbelichteten Hintergrunds. Jedoch werden bei der MSS-Anbaureihe aufgrund der starken Überbelichtung einiger Trauben teilweise 3D-Punkte von Beeren gelöscht.

Die Punktwolke wird anschließend mit dem SOR-Filter von Ausreißern bereinigt. Eine Nachbarschaftszahl von 100 Punkten und ein im Vergleich zum PHENObot strengerer Faktor $f = 1$ hat sich hier als optimal erwiesen, da aufgrund der starken Überbelichtung die Anzahl falsch triangulierter Punkte höher liegt als beim PHENObot.

Zuletzt werden die Punktwolken mit einem Glättungsradius $r_s = 5mm$ und einem Polynomgrad von zwei nach dem MLS-Verfahren geglättet.

5.1.5 Detailgrad der Punktwolken

Abbildung 5.12 stellt zwei Trauben der TS- und MSS-Anbaureihe vergleichend nebeneinander dar. Der Detailgrad ist durchgehend hoch, einzelne Beeren wurden geometrisch distinktiv rekonstruiert.

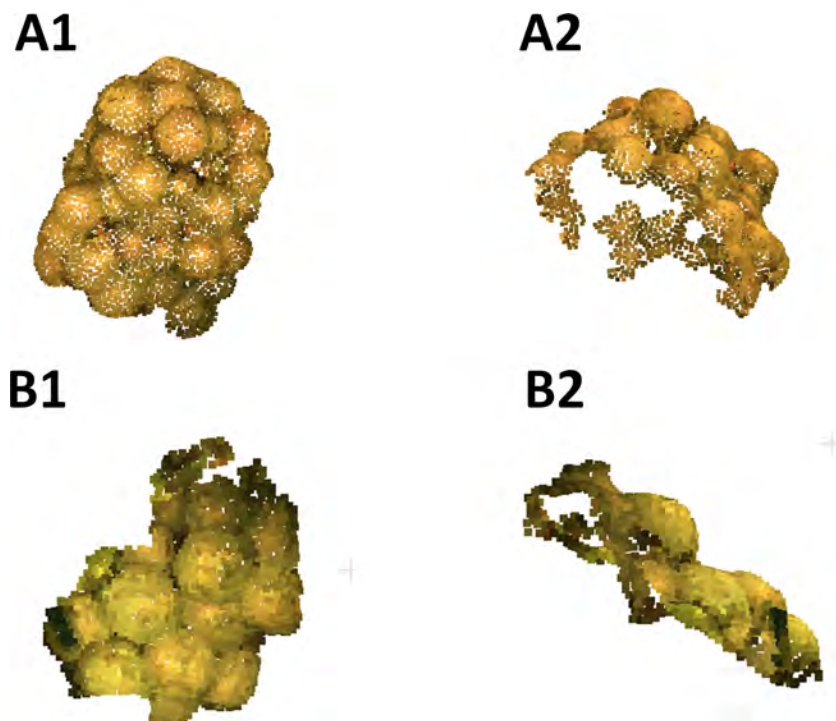


Abbildung 5.12: A1 und A2: Eine Traube der TS-Anbaureihe in frontaler bzw. seitlicher Perspektive. Individuelle Beeren wurden geometrisch distinktiv von anderen Beeren unterscheidbar und rund rekonstruiert. **B1 und B2:** Eine Traube der MSS-Anbaureihe in frontaler bzw. seitlicher Perspektive. Auch hier sind individuelle Beeren rund rekonstruiert worden und geometrisch deutlich voneinander abgrenzbar.

Der Detailgrad wird jedoch merklich reduziert, wenn die Trauben überbelichtet aufgenommen werden. Abbildung 5.13 zeigt die Auswirkungen der Überbelichtung auf die Rekonstruktion der Punktwolken und den Effekt der Hintergrundentfernung auf die Trauben. Punkte der Traube wurden fälschlicherweise dem Hintergrund zugeordnet und entfernt. Außerdem ist geometrische Detailgrad im Vergleich zu den nicht überbelichteten Trauben der MSS-Anbaureihe merklich geringer. Erhebungen individueller Beeren sind zwar erkennbar, jedoch wurde die eigentlich runde Geometrie der Beeren abgeflachter rekonstruiert. Dieses Phänomen korrespondiert mit den Beobachtungen bei der Rekonstruktion der Trauben im BBCH75 Reifestadium. In beiden Fällen entstand durch die Überbelichtung eine homogene Färbung der Beerenmitte. Bildpunkte werden darauffolgend falsch zugeordnet und entweder gar nicht oder an falscher Stelle rekonstruiert, wodurch die eigentliche Geometrie der Beeren verfälscht wird.

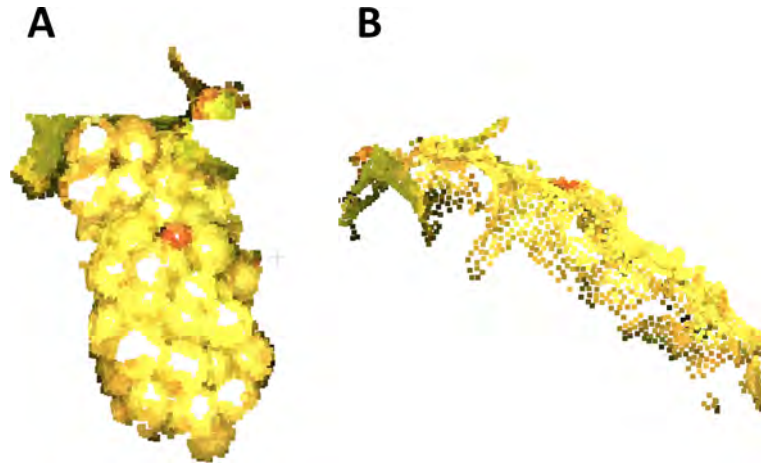


Abbildung 5.13: **A:** Eine überbelichtete Traube der MSS-Anbaureihe in frontaler Perspektive. Deutlich erkennbar sind die während der Hintergrundentfernung fälschlicherweise gelöschten 3D-Punkte in der Beerenmitte. **B:** Dieselbe Traube der MSS-Anbaureihe in seitlicher Perspektive.

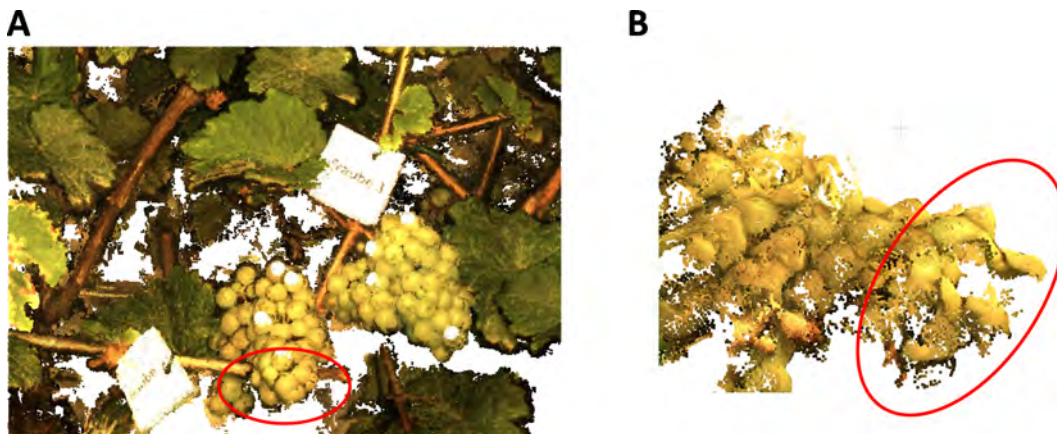


Abbildung 5.14: **A:** Der rekonstruierte Ausschnitt der Anbaureihe, der in Abbildung 5.9 in frontaler Perspektive gezeigt wird. In Rot ist die Spitze der Weintraube markiert. **B:** Die markierte Weintraube in seitlicher Perspektive. Der markierte Bereich wurde doppelt rekonstruiert. Dies ist höchstwahrscheinlich auf die Verdeckung der Kamera durch den Ast und bewegte Objektpunkte während der Bildaufnahme zurückzuführen.

Dynamische Objekte während der Bildaufnahme haben ebenfalls Einfluss auf die Rekonstruktion. Abbildung 5.14 zeigt den Effekt, den während der Bildaufnahme bewegte Objekte auf die Rekonstruktion haben. Die Abbildung zeigt den rekonstruierten Ausschnitt der MSS-Anbaureihe aus Abbildung 5.9 im vorherigen Unterkapitel. Ein Ast berührte während der Bildaufnahme den Phenoliner, bewegte sich und verdeckte zudem einen Großteil der Weinrebe. Dies führte dazu, dass Teile der Weinrebe doppelt rekonstruiert wurden. Grund

dafür ist die Dynamik der homologen Bildpunkte. Die SfM/MVS Methode geht von unbewegten Objekten aus. Wechseln als homolog identifizierte Bildpunkte während der Aufnahme ihre Position im Objektraum, wird die Schätzung der relativen Kamerapositionen zwischen den Aufnahmezeitpunkten und damit die Triangulation der Bildpunkte verfälscht.

5.1.6 Zusammenfassung

Bildaufnahme und Zeitaufwand

Die Datenerhebung mit dem Phenoliner ist durch die Abschirmung der Sensorik und der Weinreben im Tunnel und die Verwendung von Beleuchtungseinheiten unabhängig von äußeren Umweltbedingungen möglich. Das vertikale MKS und der IggGeotager 2.0 ermöglichen eine hohe bildliche Abdeckung der Anbaureihen von 95% in der Horizontalen und 75% in der Vertikalen. Die Bildaufnahme geschieht automatisiert und mit einer hohen Aufnahmefrequenz von $5Hz$, die eine Datenerhebung aus der Bewegung im Hochdurchsatz gestattet. Drei Kilometer von Anbaureihen können in 11 Stunden erfasst werden. Sensorik und Recheneinheiten sind geschützt auf dem Phenoliner montiert. Über GPS- und Inklinometerdaten werden die Bilder im Postprocessing mit Zentimetergenauigkeit georeferenziert.

Beide Erziehungssysteme können mit dem Phenoliner erfasst werden. Im Falle der TS-Anbaureihen besitzt der Anwender aufgrund der geringen Tiefe ihres Blattwerks mehr Spielraum zur Positionierung des Tunnels über den Anbaureihen. Der Abstand des MKS kann daher etwas variiert und auf eine optimale Entfernung angepasst werden. Dieser Spielraum ist beim MSS-Erziehungssystem aufgrund des großen Volumens des Blattwerks eingeschränkt. Um den Tunnel über die Anbaureihen bewegen zu können, muss der Tunnel mittig zur Anbaureihe platziert werden. Dadurch nimmt das MKS zwangsweise einen näheren Abstand zu den Anbaureihen ein. Aufgrund der geringeren Distanz und des größeren Volumens des Blattwerks kommt es daher im MSS-Erziehungssystem stellenweise zu Berührungen von herausragenden Ästen mit dem Phenoliner. Äste können dabei unter Umständen die Kameralinse verdecken.

Georeferenzierung, 3D-Rekonstruktion und Beleuchtung

Die Georeferenzierung der Bilder erreicht durch die Berücksichtigung der Fahrzeugneigung und des Hebelarms zwischen MKS und GPS-Antenne Residuen von $2 - 3cm$ in Lage und Höhe. Dies entspricht der spezifizierten Unsicherheit der RTK-GPS Positionsbestimmung. Es kann daher konstatiert werden, dass die Übertragung der Koordinate der GPS-Antenne über Berücksichtigung des Hebelarms und der Drehwinkel des Phenoliners auf die Position der Kameras durch den IggGeotager 2.0 gut funktioniert und die GPS-Positionen als verlässliche Näherungswerte für die Bündelausgleichung genutzt können. Die Ergebnisse stellen eine deutliche Verbesserung gegenüber der Georeferenzierung der Bilder des PHENObots dar, bei denen in der Höhe Residuen bis zu $30cm$ entstanden.

Die Anzahl der Bilder und damit der Perspektiven auf die Anbaureihen ist deutlich höher als im Falle des PHENObots. Objekte werden daher mit höherer Wahrscheinlichkeit rekonstruiert, da eine ausreichend hohe Menge von Beobachtungen genutzt werden kann. Der durchschnittliche Rückprojektionsfehler der 3D-Rekonstruktion ist jedoch höher, als im Falle des PHENObots. Dies könnte mit der kürzeren horizontalen Entfernung $base_h = 3,3cm$ zwischen den Bildern zusammenhängen, welche schleifende Schnittwinkel zwischen den Projektionsstrahlen verursacht. Künftige Experimente sollten eine optimale Entfernung zwischen den Bildern bestimmen, die eine hohe Bildabdeckung bei einem möglichst geringen Rückprojektionsfehler ermöglicht.

Die sechs Halogenstrahler erhellen die Anbaureihen alle Bilder und erzeugen eine warme Farbtemperatur. Die Bilder besitzen daher einen Gelbstich. Dieser Umstand ist jedoch unerheblich für die Datenklassifikation ist, da relative farbliche Unterschiede der Pflanzenteile gewahrt bleiben. In der TS-Anbaureihe können einige wenige Fälle von Überbelichtung auf der Beerenmitte beobachtet werden. Vorallem im Falle der MSS-Anbaureihen erweisen sich sechs Strahler jedoch als überproportioniert. Aufgrund des ohnehin geringeren Abstands zur Anbaureihe, der großen Tiefe des Blattwerks sowie der starken Lichtquellen sind Trauben der MSS-Anbaureihe deshalb deutlich häufiger überbelichtet.

Die oben geschilderten Phänomene traten beim PHENObot nicht auf, da aufgrund seiner geringen Größe und des technisch fortschrittlicheren Beleuchtungsrahmens eine optimale Beleuchtung und Distanz zu den Weinreben erreicht werden konnte.

Die runde Form einzelner Beeren wurde mit hohem geometrischen Detailgrad rekonstruiert, so dass einzelne Beeren geometrisch distinktiv voneinander unterscheidbar sind. Die geometrische Detailgrad der Trauben der MSS-Anbaureihe ist jedoch aufgrund der Überbelichtung und der daraus resultierenden homogenen Färbung merklich geringer. Tiefer im Blattwerk liegende und daher nicht überbelichtete Trauben konnten besser rekonstruiert werden. Durch die homogene Färbung der Trauben wird die Geometrie der Beeren daher teilweise nicht ausreichend exakt rekonstruiert, was die korrekte Detektion von Beeren erschwert. Hier sind weitere Experimente zu den Beleuchtungseinheiten notwendig, um eine gleichmäßige Ausleuchtung der Anbaureihe ohne Überbelichtung und direkte Reflexionen auf der Beerenmitte zu erreichen. Zusätzlich können die Verdeckung der Linse und bewegte Objekte ebenfalls zu einer korruptierten Rekonstruktion führen.

Das Prinzip eines homogen gefärbten Hintergrundes erweist sich grundsätzlich für die Unterscheidung zwischen Vorder- und Hintergrund als geeignet. Die Wahl eines weißen Hintergrundes ist aufgrund des verwendeten Materials und der starken Reflexionseigenschaften jedoch nicht optimal. Der Stoff des Tuches bewegt sich durch den Fahrtwind und wirft mitunter Schatten. Das Licht der Halogenstrahler wird in einem hellgelben Farbton reflektiert und erzeugt einen Farbton ähnlich zu dem der Trauben. Die Unterscheidung zwischen Vorder- und Hintergrund wird dadurch erschwert.

5.2 Datenklassifikation

5.2.1 Ermittlung optimaler Radien für den SFHc-Deskriptor

Da sich Sensorplattform und Belichtung von den Voruntersuchungen mit dem PHENObot unterscheiden, können nicht ohne weiteres die in den Voruntersuchungen für das BBCH89 Reifestadium bestimmten, optimalen Radien für r_n und r_H des SFHc-Deskriptors für die Klassifikation der Punktwolken aus Phenoliner Bilddaten verwendet werden. Für die Klassifikation der Punktwolken aus den Bildern des Phenoliners müssen daher erneut optimale Radien r_n und r_H für den SFHc-Deskriptor bestimmt werden. Dabei wird wie in Kapitel 4.2 und gemäß Abbildung 4.11 vorgegangen. Da die SFHc-Deskriptoren in diesem Kapitel bereits als bester Deskriptor für die Klassifikation von Weinreben identifiziert wurden, werden andere Deskriptoren für die Klassifikation der Phenoliner Punktwolken nicht abermalig untersucht.

Für r_n werden Werte im Bereich $[1mm..5mm]$ mit $1mm$ Schrittweite, für den Histogrammradius r_H Werte im Bereich $[2..17mm]$ mit $1mm$ Schrittweite verwendet. Die Suche nach optimalen Radien wird für jedes Erziehungssystem separat durchgeführt, um die farblichen und geometrischen Unterschiede zwischen beiden Erziehungssystemen zu berücksichtigen, welche bei der Rekonstruktion durch die Überbelichtung entstanden. Entsprechend der Ergebnisse aus Kapitel 4.2.1 wird die Klassifikation inklusive der Klassenglättung durchgeführt.

Zur Untersuchung aller Radienkombinationen werden aus den rekonstruierten Anbaureihen Testdaten P^t mit manuell zugewiesenen Klassen $c = 1, 2$, sowie Trainingsdaten $T \notin P^t$ erstellt und die optimalen Radien anhand von Accuracy, Recall und Precision berechnet. Für das TS- und MSS-Erziehungssystem werden jeweils eigene Trainings- und Testdaten erstellt. Die Eignung jeder möglichen Kombination von r_n und r_H wird zuletzt hinsichtlich der mit ihr erreichbaren Accuracy, Recall und Precision bewertet. Tabelle 5.2 listet die ermittelten optimalen Radien, sowie Recall und Precision für die Testdaten P^t der TS- und MSS-Anbaureihe auf.

Für die TS-Anbaureihe ergibt sich die gleiche Radienkombination, die in Kapitel 4 für das Reifestadium BBCH89 bestimmt wurde. Damit wird die in den Voruntersuchungen ermittelte optimale Radienkombination für den SFHc-Deskriptor nochmals bestätigt.

	r_n [mm]	r_H [mm]	Recall[%]	Precision[%]
TS 89	3	9	96	76
MSS 89	3	5	89	32

Tabelle 5.2: Optimale Radien, Recall und Precision für P^t aus den Phenoliner Aufnahmen.

Für die MSS-Anbaureihe hingegen ergab sich im Vergleich zu diesen Voruntersuchungen zwar ein identischer Radius $r_n = 3mm$, jedoch ein davon abweichender Radius $r_H = 5mm$ (PHENObot $r_H = 9mm$). Der unterschiedliche Radius resultiert dabei vermutlich aus der stellenweisen Überbelichtung der Bilder und dem dadurch verursachten farblichen Unterschied sowie der qualitativ schlechteren Rekonstruktion der Weinreben, die sich bereits an der höheren Varianz des Rückprojektionsfehlers der MSS-Anbaureihe im Vergleich zur TS-Anbaureihe erkennen lässt (Tabelle 5.1).

Das Ergebnis demonstriert abermals, dass optimale Berechnungsparameter für SFHc-Deskriptoren individuell für Variationen der Morphologie der Weinreben durch Belichtung, Reifestadium, Erziehungssystem oder Sorte bestimmt werden müssen, um optimale Ergebnisse bei der Klassifikation zu erzielen.

5.2.2 Klassifikation der fünf Abschnitte beider Erziehungssysteme

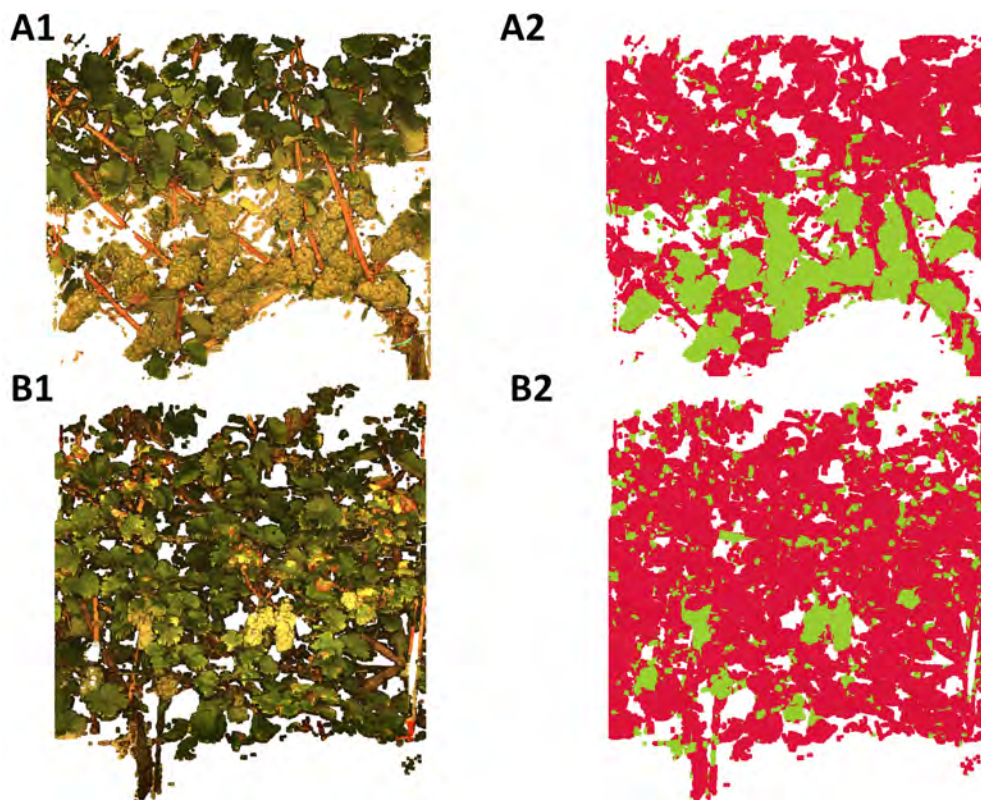


Abbildung 5.15: **A1:** Ein vorverarbeiteter 1-Meter Abschnitt der TS-Anbaureihe. **A2:** Der klassifizierte Abschnitt der TS-Anbaureihe. Rot: Rest. Grün: Trauben. **B1:** Ein vorverarbeiteter 1-Meter Abschnitt der MSS-Anbaureihe. **B2:** Der klassifizierte Abschnitt der MSS-Anbaureihe. Rot: Rest. Grün: Trauben.

Die Klassifikation wird mit den ermittelten optimalen Radien ausschließlich für die fünf Abschnitte der beiden Anbaureihen durchgeführt, die im nächsten Kapitel die Referenzwerte für die Schätzung der Ertragsparameter und des Ertrages liefern (s. 5.1.2).

Abbildung 5.15 zeigt beispielhaft zwei Abschnitte aus der TS- bzw. MSS-Anbaureihe (nicht die Testdaten) und das Ergebnis ihrer Klassifikation mit optimalen Radien des SFHc-Deskriptors.

Der erreichte Recall und Precision der Klassifikation der P^t des TS- und MSS-Erziehungssystems (Tabelle 5.2) sind ähnlich zu denen aus der Klassifikation der Anbaureihen im BBCH89 Reifestadium, die mit dem PHENObot erfasst wurden (TS: Recall 94%, Precision 63%; MSS: Recall 89%, Precision 39%). Aus den Ergebnissen der Klassifikation der Testdaten P^t , sowie der visuellen Inspektion der Klassifikation der einzelnen Abschnitte und den bisherig gemachten Erfahrungen mit den Daten des PHENObots kann davon ausgegangen werden, dass

- die Klassifikation der fünf Abschnitte anhand der ermittelten optimalen Radien ähnliche Werte für Recall und Precision erreicht, wie in den Voruntersuchungen und
- dass die Klassifikationsstrategie auch für Punktwolken geeignet ist, die aus Bilddaten des Phenoliners rekonstruiert worden sind.

Individuelle Ergebnisse für Recall und Precision der Klassifikation der einzelnen 1m-Abschnitte wurden daher nicht erhoben. Der Großteil der Trauben abbildenden 3D-Punkte wird detektiert. In den klassifizierten Daten existieren Punktregionen, die tatsächlich zum Hintergrund oder zu Blättern oder Ästen gehören, aber stellenweise der Klasse **Trauben** zugewiesen wurden. Im Falle der TS-Anbaureihe wurden einige wenige Trauben teilweise vollständig falsch als Punkte des Hintergrunds klassifiziert. Die ist höchstwahrscheinlich auf die farbliche Ähnlichkeit von Vorder- und Hintergrund zurückzuführen.

Die Punkte der Klasse **Trauben** der fünf Abschnitte werden im folgenden Kapitel für die Schätzung der Ertragsparameter und des Ertrages genutzt.

5.3 Quantifikation: Ertragsschätzung aus Ertragsparametern

Wie in Kapitel 4.3.2 der Voruntersuchungen bereits beschrieben², setzen sich die tatsächlichen Gesamtanzahlen der Beeren und Trauben den obigen Überlegungen folgend aus den vier Teilanzahlen $\mathcal{D}^{B/T}$, $\mathcal{X}^{B/T}$, $\mathcal{M}^{B/T}$ und $\mathcal{O}^{B/T}$ zusammen. Gemäß Formel 4.15 ergibt sich die Gesamtanzahl aus der Summe der vier Teilanzahlen.

Die Voruntersuchungen zeigten, dass die Precision von $\mathcal{D}^{B/T}$ für das BBCH89 Reifestadium sowohl für die Beeren- als auch die Traubenzählung sehr hoch ist (mind. 95 – 100%, vgl. Tabellen 4.11 und 4.14). $\mathcal{X}^{B/T}$ ist folglich klein und kann vernachlässigt werden. Die Situation der Schätzung der Gesamtanzahlen gestaltet sich demnach für die Beeren und Trauben folgendermaßen:

- Die Gesamtanzahl der Trauben setzt sich aus der Anzahl detektierter sichtbarer Trauben $\mathcal{D}^T$, der Anzahl von der Auswertemethode entfernten Trauben $\mathcal{M}^T$ sowie der Anzahl der stets unsichtbaren Trauben $\mathcal{O}^T$ zusammen.
- Die geschätzte Gesamtanzahl der Beeren besteht zum einen Teil aus den detektierten Beeren der detektierten Trauben $\mathcal{D}^T$. Pro detektierte Traube müssen noch die Anzahlen der Beeren hinzugezählt werden, die 1) nicht detektiert wurden ($\mathcal{M}^B$) und 2) die im Inneren und auf der Rückseite der Trauben liegen und daher nicht gesehen werden können ($\mathcal{O}^B$).
Der andere Teil besteht aus den nicht detektierbaren Beeren der $\mathcal{M}^T$ und $\mathcal{O}^T$. Die Beeren der $\mathcal{M}^T$ seien entsprechend ebenfalls mit $\mathcal{M}^B$, die der $\mathcal{O}^T$ mit $\mathcal{O}^B$ bezeichnet. Für eine Schätzung Gesamtanzahl der Beeren ist somit die vorherige Schätzung von $\mathcal{M}^T$ und $\mathcal{O}^T$ nötig.

²In Kapitel 4.3.2 wurden die Teilanzahlen $\mathcal{D}^{(B/T)}$ der sichtbaren und detektierten Beeren bzw. Trauben evaluiert. Die dort vorgestellten Ergebnisse für Recall und Precision beziehen sich also auf $\mathcal{V}^{(B/T)}$ - nicht hingegen auf die Gesamtanzahl der Beeren und Trauben.

Um die Gesamtanzahlen zu erhalten, müssen $\mathcal{M}^{B/T}$ und $\mathcal{O}^{B/T}$ geschätzt werden. Dazu müssen Kalibrierwerte abgeleitet werden, mit deren Hilfe aus $\mathcal{D}^{B/T}$ Schätzungen $\hat{\mathcal{M}}^B$ bzw. $\hat{\mathcal{M}}^T$ und $\hat{\mathcal{O}}^B$ bzw. $\hat{\mathcal{O}}^T$ bestimmt werden können. Die Gesamtanzahl wird Formel 4.15 folgend über:

$$\hat{\mathcal{E}}^{B/T} = \mathcal{D}^{B/T} + \hat{\mathcal{M}}^{B/T} + \hat{\mathcal{O}}^{B/T}, \quad (5.13)$$

geschätzt, wobei $\hat{\mathcal{E}}^{B/T}$ entsprechend die geschätzte Gesamtzahl $\hat{\mathcal{E}}^B$ bzw. $\hat{\mathcal{E}}^T$ (estimated) der Trauben und Beeren ist. Abbildung 5.16 illustriert die Situation bei der Schätzung der Gesamtanzahlen.

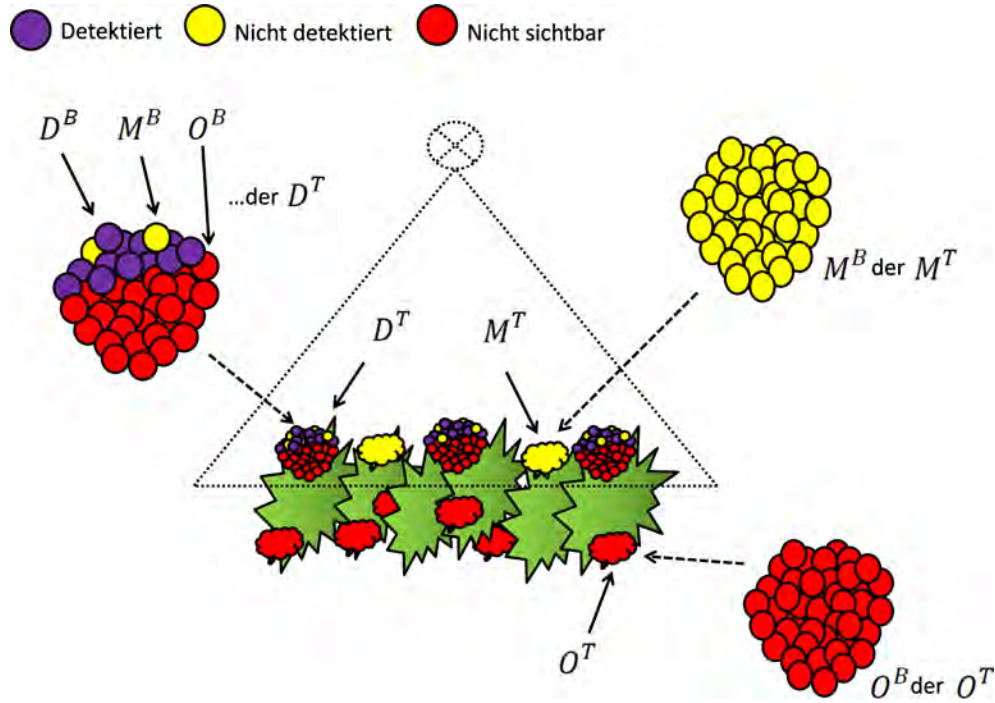


Abbildung 5.16: In der Bildmitte die Darstellung von $\mathcal{D}^T$, $\mathcal{M}^T$ und $\mathcal{O}^T$ in einer Anbaureihe aus der Vogelperspektive. Das schwarz gestrichelte Dreieck stellt das Sichtfeld der Kamera dar. Gestrichelte Pfeile: Detaillierte Darstellung der Beerenzählung für $\mathcal{D}^T$, $\mathcal{M}^T$ und $\mathcal{O}^T$.

In diesem Kapitel werden zwei Ziele verfolgt:

1. Die Kalibrierwerte, mit deren Hilfe aus $\mathcal{D}^B$ bzw. $\mathcal{D}^T$ die Gesamtanzahl der Beeren und Trauben $\hat{\mathcal{E}}^B$ bzw. $\hat{\mathcal{E}}^T$ geschätzt werden kann, sollen bestimmt werden. Auch der durch das Connected Component-Verfahren verursachte Fehler, bei dem ggf. mehrere Trauben in einer Traube vereinigt werden, soll durch eine entsprechende Kalibrierung berücksichtigt werden (s. Kapitel 4.3.2). Zur Berechnung und Überprüfung der Kalibrierwerte werden Referenzwerte für die einzelnen Anzahlen erhoben.
2. Aus den so geschätzten Gesamtanzahlen soll anschließend der Gesamtertrag geschätzt werden. Auch hierfür werden Kalibrierwerte benötigt, welche die Gesamtanzahlen in Bezug zu ihrem Gewicht setzen. Der geschätzte Gesamtertrag sei mit $\hat{\mathcal{Y}}$ (yield) bezeichnet und wird als Gewicht ausgedrückt. Es wird also eine Funktion benötigt, die aus der Anzahl der Beeren und der Trauben den Ertrag schätzt:

$$\hat{\mathcal{Y}} = f(\hat{\mathcal{E}}^B, \hat{\mathcal{E}}^T). \quad (5.14)$$

Zunächst wird ein Überblick über die allgemeine Vorgehensweise zur Ertragsschätzung, sowie eine tabellarische Übersicht über die Kalibrier-, Schätz- und Referenzwerte gegeben. Anschließend wird die Berechnung der Kalibrierwerte aus den Referenzwerten, sowie eine Überprüfung der mit den Kalibrierwerten erzielten

Ergebnisse anhand der Referenzwerte geschildert. Mit den Kalibrierwerten werden darauffolgend für die fünf markierten Abschnitte jedes Erziehungssystems (s. Kapitel 5.1.2) die Ertragsparameter bestimmt und Ertragsschätzungen gemäß Formel 5.14 vorgenommen. Dabei wird der Ertrag für jeden der fünf Abschnitte sowie der summierte Ertrag aus allen Abschnitten geschätzt und mit jeweiligen Referenzwerten für den Abschnitt verglichen. Eine abschließende Diskussion der Ergebnisse der Ertragsschätzung erfolgt im Kapitel 5.3.6.

5.3.1 Ablauf der Schätzung der Ertragsparameter und des Ertrages

Abbildung 5.17 stellt den Ablauf der Schätzung der Gesamtanzahlen und des Gesamtertrages anhand der Kalibrierwerte in einem Flussdiagramm dar. Anhand der Kalibrierwerte $\hat{M}_u^T$, f_m und f_o werden zunächst die Gesamtanzahlen der Trauben geschätzt (Schritte 1-7). Anschließend erfolgt die Schätzung der Gesamtanzahlen der Beeren der detektierten Trauben über f_B und der Beeren der entfernten und stets verdeckten Trauben über B_T (Schritte 8-12). Der Ertrag wird anschließend durch Summierung der geschätzten Beeren und Multiplikation mit dem Kalibrierwert g_B geschätzt (Schritte 13-16).

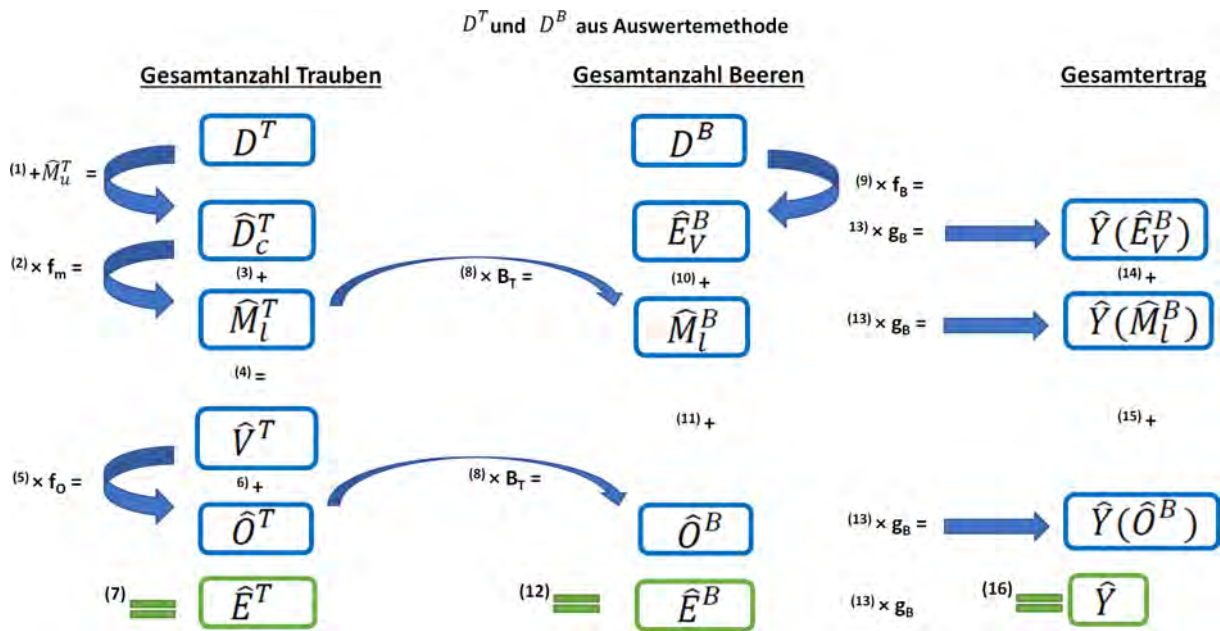


Abbildung 5.17: Die Berechnung der Gesamtanzahlen der Beeren und Trauben basiert auf der Anzahl detektierter Beeren und Trauben D^B bzw. D^T . Unter Nutzung der geschätzten Teil- und Gesamtanzahlen wird der Gesamtertrag des betrachteten Abschnitts geschätzt. Die Reihenfolge der Berechnungsschritte und der erhaltenen Ergebnisse ist entsprechend nummeriert. Mit Pfeilen: Multiplikation einer Ausgangsgröße mit einem Kalibrierwert. Der Pfeil weist auf ein sich dadurch ergebendes Zwischenergebnis hin. Ohne Pfeile: Addition der Zwischenergebnisse.

Zur Bestimmung der Kalibrierwerte werden Referenzwerte benötigt:

- Referenzwerte für die Anzahl der Trauben werden in den fünf markierten Abschnitten jedes Erziehungssystems jeweils pro Abschnitt erhoben. Für die TS-Anbaureihe ergaben sich insgesamt 133, für die MSS-Anbaureihe 277 Trauben.
- Referenzwerte für die Anzahl der Beeren ergeben sich aus den 25 fortlaufend nummerierten Referenztrauben. Pro Abschnitt wurden je fünf Trauben markiert.
- Referenzwerte für das Gewicht einzelner Trauben ergeben sich durch Wiegen der 25 Referenztrauben. Referenzwerte für das Gewicht der Trauben pro Abschnitt ergeben sich durch Wiegen aller Trauben des Abschnitts.

Die Tabellen 5.3, 5.4 und 5.5 fassen alle genutzten Variablen geordnet nach ihrem Bezug zu Beeren, Trauben und Ertrag zusammen. Die meisten Variablen werden bereits in Abbildung 5.17 verwendet, einige werden erst im weiteren Verlauf der Arbeit gebraucht. Die dritte Spalte beschreibt die Funktion der Variable, die vierte Spalte ordnet sie den Typen Kalibrier, Referenz-, Zählungs- und Schätzwert zu, die letzte Spalte erläutert die Herkunft bzw. Formel zur Berechnung des Wertes.

Bezug	Variable	Beschreibung	Typ	Formel/Herkunft
Trauben	$\mathcal{N}^T$	Gesamtanzahl aller Trauben.	R	Manuell in Bildern oder bei der Ernte gezählt
Trauben	$\mathcal{D}^T$	Anzahl detektierter Trauben.	Z	Über Auswertemethode
Trauben	$\mathcal{M}_u^T$	Anzahl der während des CC-Verfahrens in einer Komponente vereinten Trauben.	R	Manuell gezählt
Trauben	$\mathcal{M}_l^T$	Anzahl der entfernten Trauben.	R	Manuell gezählt
Trauben	$\mathcal{M}^T$	Gesamtanzahl sichtbarer, aber nicht detektierter Trauben.	R	$\mathcal{M}^T = \mathcal{M}_u^T + \mathcal{M}_l^T$ (5.15)
Trauben	$\hat{\mathcal{M}}_u^T$	Geschätzter additiver Wert, der $\mathcal{M}_u^T$ approximiert. Korrigiert $\mathcal{D}^T$ um die Anzahl der Trauben, die in einer Komponente vereint sind.	K	$\hat{\mathcal{M}}_u^T = \frac{\mathcal{D}^B}{m_B}$ (5.16)
Trauben	$\hat{\mathcal{D}}_c^T$	Um $\hat{\mathcal{M}}_u^T$ korrigierte Anzahl detektierter Trauben.	S	$\hat{\mathcal{D}}_c^T = \mathcal{D}^T + \hat{\mathcal{M}}_u^T$ (5.17)
Trauben	$\mathcal{V}^T$	Anzahl sichtbarer Trauben.	R	Manuell in Bildern gezählt
Trauben	f_m	Kalibrierwert zur Schätzung der Anzahl sichtbarer, aber nicht detektierter Trauben $\mathcal{M}_l^T$.	K	$f_m = \frac{\mathcal{M}_l^T}{\hat{\mathcal{D}}_c^T}$ (5.18)
Trauben	$\hat{\mathcal{M}}_l^T$	Geschätzte Anzahl nicht detektierter Trauben.	S	$\hat{\mathcal{M}}_l^T = \hat{\mathcal{D}}_c^T \cdot f_m$ (5.19)
Trauben	$\mathcal{O}^T$	Anzahl stets verdeckter Trauben.	R	Manuell in Bildern gezählt oder nach Formel 4.14
Trauben	$\hat{\mathcal{V}}^T$	Geschätzte Anzahl sichtbarer Trauben.	S	$\hat{\mathcal{V}}^T = \hat{\mathcal{D}}_c^T + \hat{\mathcal{M}}_l^T$ (5.20)
Trauben	f_o	Kalibrierwert zur Schätzung der Anzahl stets verdeckter Trauben aus $\hat{\mathcal{V}}^T$.	K	$f_o = \frac{\mathcal{O}^T}{\hat{\mathcal{V}}^T}$ (5.21)
Trauben	$\hat{\mathcal{O}}^T$	Geschätzte Anzahl stets verdeckter Trauben.	S	$\hat{\mathcal{O}}^T = \hat{\mathcal{V}}^T \cdot f_o$ (5.22)
Trauben	$\hat{\mathcal{E}}^T$	Ertragsparameter: Geschätzte Gesamtanzahl aller Trauben. Summe der detektierten, nicht detektierten und stets verdeckten Trauben.	S	$\hat{\mathcal{E}}^T = \hat{\mathcal{D}}_c^T + \hat{\mathcal{M}}_l^T + \hat{\mathcal{O}}^T$ (5.23)

Tabelle 5.3: Variablen mit Bezug zu Trauben. K: Kalibrierwert, Z: Zählung R: Referenzwert, S: Schätzung

Bezug	Variable	Beschreibung	Typ	Formel/Herkunft
Beeren	$\mathcal{N}^B$	Gesamtanzahl aller Beeren pro Referenztraube.	R	Manuell gezählt
Beeren	$\mathcal{D}^B$	Anzahl detektierter Beeren der detektierten Trauben.	Z	Über Auswertemethode
Beeren	B_T	Durchschnittliche Gesamtanzahl der Beeren pro Traube, mit der aus $\hat{\mathcal{M}}_l^T$ und $\hat{\mathcal{O}}^T$ die Anzahl ihrer Beeren $\hat{\mathcal{M}}_l^B$ und $\hat{\mathcal{O}}^B$ geschätzt wird.	K	Empirisch ermittelt
Beeren	f_B	Korrigiert $\mathcal{D}^B$ der detektierten Trauben um die Anzahl ihrer nicht detektierten und nicht sichtbaren Beeren.	K	$f_B = \frac{\mathcal{N}^B}{\mathcal{D}^B}$ (5.24)
Beeren	g_B	Durchschnittliches Gewicht einer Beere.	K	Empirisch ermittelt
Beeren	m_B	Durchschnittliche Anzahl detektierter Beeren pro detektierte Traube.	K	Empirisch ermittelt
Beeren	$\hat{\mathcal{E}}_V^B$	Geschätzte Gesamtanzahl der Beeren der detektierten Trauben.	S	$\hat{\mathcal{E}}_V^B = \mathcal{D}^B \cdot f_B$ (5.25)
Beeren	$\hat{\mathcal{M}}_l^B$	Geschätzte Anzahl der Beeren der nicht detektierten Trauben.	S	$\hat{\mathcal{M}}_l^B = \hat{\mathcal{M}}_l^T \cdot B_T$ (5.26)
Beeren	$\hat{\mathcal{O}}^B$	Geschätzte Anzahl der Beeren der stets verdeckten Trauben.	S	$\hat{\mathcal{O}}^B = \hat{\mathcal{O}}^T \cdot B_T$ (5.27)
Beeren	$\hat{\mathcal{E}}^B$	Ertragsparameter: Geschätzte Gesamtanzahl aller Beeren aller Trauben.	S	$\hat{\mathcal{E}}^B = \hat{\mathcal{E}}_V^B + \hat{\mathcal{M}}_l^B + \hat{\mathcal{O}}^B$ (5.28)

Tabelle 5.4: Variablen mit Bezug zu Beeren. K: Kalibrierwert, Z: Zählung R: Referenzwert, S: Schätzung

Bezug	Variable	Beschreibung	Typ	Formel/Herkunft
Ertrag	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$	Geschätztes Gewicht der detektierten Trauben über die geschätzte Anzahl ihrer Beeren.	S	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B) = \hat{\mathcal{E}}_V^B \cdot g_B$ (5.29)
Ertrag	$\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)$	Geschätztes Gewicht der nicht detektierten Trauben über die geschätzte Anzahl ihrer Beeren.	S	$\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B) = \hat{\mathcal{M}}_l^B \cdot g_B$ (5.30)
Ertrag	$\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B)$	Geschätztes Gewicht der stets verdeckten Trauben über die geschätzte Anzahl ihrer Beeren.	S	$\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B) = \hat{\mathcal{O}}^B \cdot g_B$ (5.31)
Ertrag	$\mathcal{Y}^T$	Gewicht einer Referenztraube.	R	Gewogen
Ertrag	$\mathcal{Y}^A$	Ertrag eines Abschnitts.	R	Gewogen
Ertrag	$\hat{\mathcal{Y}}$	Geschätzter Gesamtertrag.	S	$\hat{\mathcal{Y}} = \hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B) + \hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B) + \hat{\mathcal{Y}}(\hat{\mathcal{O}}^B)$ (5.32)

Tabelle 5.5: Variablen mit Bezug zum Ertrag. K: Kalibrierwert, R: Referenzwert, S: Schätzung

5.3.2 Berechnung der Kalibrierwerte

Für jeden der fünf Abschnitte der TS- und MSS-Anbaureihe werden zunächst $\mathcal{D}^B$ und $\mathcal{D}^T$ mit der Auswertemethode bestimmt. Da Sorte und Reifestadium identisch zu den Untersuchungen mit dem PHENObot sind (s. Kapitel 4.3.2.3), werden für das CC-Verfahren und den Berryfinder die identischen Parameter genutzt. Um von der Berechnung der Kalibrierwerte unabhängige Größen zu erhalten, werden die Kalibrierwerte ausschließlich anhand der Referenzwerte der ersten drei Abschnitte berechnet. Auf diese Weise verbleiben die

zwei letzten Abschnitte als unabhängige Referenzwerte, anhand derer die Gültigkeit der Kalibrierwerte und der Ertragsschätzung überprüft werden kann.

5.3.2.1 Kalibrierwerte zur Schätzung der Anzahl und des Gewichts der Beeren der detektierten Trauben

Zunächst werden die Kalibrierwerte B_T , m_B , g_B sowie f_B berechnet. Dazu werden aus den Traubenkandidaten manuell die 25 rekonstruierten Referenztrauben ausgewählt und die Werte für $\mathcal{D}^B$, $\mathcal{N}^B$ sowie das Gewicht $\mathcal{Y}^T$ pro Referenztraube herangezogen. Um die Kalibrierwerte zu berechnen, werden die ersten 12 Referenztrauben genutzt. Diese befinden sich entsprechend der fortlaufenden Nummerierung der Referenztrauben im ersten, zweiten und dritten Abschnitt.

Um von Ausreißern unbeeinflusste Schätzungen zu erhalten, wird aus allen Werten der Median als Kalibrierwert ausgewählt. B_T und m_B sind jeweils der Median von $\mathcal{N}^B$ bzw. $\mathcal{D}^B$. Das Gewicht einer Einzelbeere g_B wird berechnet, indem das tatsächliche Gewicht $\mathcal{Y}^T$ einer Referenztraube durch die tatsächliche Beerenanzahl $\mathcal{N}^B$ dividiert wird. Der Kalibrierwert f_B ist das Verhältnis von $\mathcal{N}^B$ und $\mathcal{D}^B$. Er wird gemäß Formel 5.24, Tabelle 5.4 für jede Referenztraube berechnet.

Es bleibt festzustellen, ob der ermittelte Wert für f_B repräsentativ für alle Trauben und inwieweit anhand $\hat{\mathcal{E}}_V^B$ und g_B eine Ertragsschätzung der Trauben möglich ist. Dies wird sowohl von der natürlichen Varianz der Beerenanzahl pro Traube, der Vollständigkeit der rekonstruierten Traube und der Varianz des Gewichts einer einzelnen Beere abhängen. Dazu werden die von den Kalibrierwerten unabhängigen 13 verbleibenden Referenztrauben genutzt.

Berechnung der Kalibrierwerte für die TS-Anbaureihe

Abbildung 5.18 zeigt eine rekonstruierte Referenztraube der TS-Anbaureihe und die mit dem Berryfinder detektierten Beeren. Eine visuelle Überprüfung zeigt, dass der Großteil der Beeren detektiert wurde und die Kugeln mit tatsächlichen Beeren korrespondieren.

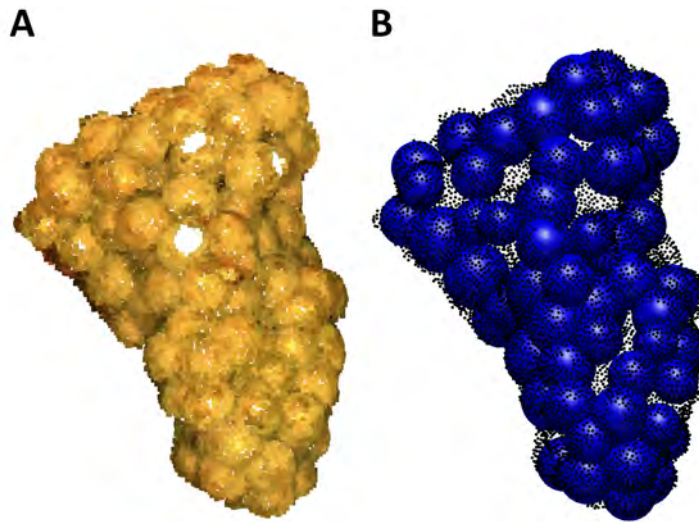


Abbildung 5.18: **A:** Eine rekonstruierte Referenztraube der TS-Anbaureihe. **B:** Mit dem Berryfinder detektierte Beeren in der Referenztraube. Kugeln können trotz fehlender Punkte auf der Beerenmitte durch die Randpunkte der Beere approximiert werden.

Beeren, denen 3D-Punkte fehlen, können durch Kugeln approximiert werden, wenn noch genügend 3D-Punkte ihre Oberfläche abbilden. Tabelle 5.6 listet die einzelnen Ergebnisse der Kalibrierwerte für jede der 12 Referenztrauben auf.

	$\mathcal{D}^B$	$\mathcal{N}^B$	f_B	$\mathcal{Y}^T[g]$	$g_B[g]$
	pro Traube	pro Traube	pro Traube	pro Traube	pro Beere
Traube 1	60	153	2,55	319,4	2,1
Traube 2	49	177	3,61	295,7	1,6
Traube 3	42	150	3,57	213,3	1,4
Traube 4	42	169	4,02	253,5	1,5
Traube 5	39	147	3,77	182,3	1,2
Traube 6	50	159	3,18	255,3	1,6
Traube 7	32	96	3,00	157,2	1,6
Traube 8	60	163	2,72	301,9	1,8
Traube 9	43	99	2,30	175,0	1,7
Traube 10	29	73	2,52	133,2	1,8
Traube 11	15	139	9,27	177,6	1,2
Traube 12	65	155	2,38	216,4	1,4
$\tilde{\text{Traube}}$	$m_B = 43$	$B_T = 151$	$\widetilde{f}_B = 3,09$		$g_B = 1,6$

Tabelle 5.6: Berechnung des Kalibrierwertes f_B , sowie des Medians der detektierten Beerenanzahl m_B , der tatsächlichen Beerenanzahl B_T und des durchschnittlichen Gewichts einer Beere g_B aus den Referenzwerten der TS-Anbaureihe. Werte sind gerundet dargestellt.

Der Median der durchschnittlich detektierten, sichtbaren Beeren ergibt sich zu $m_B = 43$. Die durchschnittliche tatsächliche Anzahl der Beeren an einer Traube B_T ergibt sich zu $B_T = 151$. Das Gewicht einer Einzelbeere g_B ergibt sich zu $g_B = 1,6g^3$. Der Median des Kalibrierwertes f_B ergibt sich zu $\widetilde{f}_B = 3,09$. Er wird für die weiteren Berechnungen in den TS-Anbaureihen als Kalibrierwert genutzt, wobei die Notation $f_B = \widetilde{f}_B$ verwendet wird.

	$\mathcal{N}^B$	$\hat{\mathcal{E}}_V^B$	$\Delta[\%]$
	pro Traube	pro Traube	
Traube 13	170	204	20
Traube 14	54	90	66
Traube 15	70	87	24
Traube 16	164	188	15
Traube 17	182	121	-34
Traube 18	171	179	5
Traube 19	151	148	-2
Traube 20	161	167	4
Traube 21	100	102	2
Traube 22	167	167	0
Traube 23	179	192	7
Traube 24	178	216	22
Traube 25	150	161	7
$\overline{\text{Traube}}$			10
σ			23
$\sum_{i=13}^{25} \text{Traube}_i$	1897	2021	7

Tabelle 5.7: Tatsächliche Beerenanzahl $\mathcal{N}^B$, geschätzte Beerenanzahl $\hat{\mathcal{E}}_V^B$ pro Traube, ihre prozentuale Abweichung, die Abweichung ihrer Summen, die mittlere Abweichung und die Standardabweichung für das TS-Erziehungssystem.

Es wird überprüft, inwieweit f_B für alle Trauben repräsentativ ist. Dazu wird $\hat{\mathcal{E}}_V^B$ der verbleibenden, von f_B unabhängigen 13 Referenztrauben nach Formel 5.25 berechnet und mit den $\mathcal{N}^B$ der 13 Referenztrauben

³Da das Stielgerüst mitgewogen wurde, ist g_B leicht zu hoch. Der Wert wird daher leicht auf $g_B = 1,4g$ reduziert.

verglichen. Tabelle 5.7 listet die Beerenanzahlen und ihre jeweilige prozentuale Abweichung, die Abweichung ihrer Summen, die mittlere Abweichung und die Standardabweichung auf. Die Standardabweichung der Beerenzählung einer einzelnen Traube beträgt 23% bei einer mittleren Abweichung von 10%. Die Summe von $\hat{\mathcal{E}}_V^B$ wird um 7% größer als die Summe $\mathcal{N}^B$ geschätzt.

Der Ertrag $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ jeder einzelnen der 13 Referenztraube wird gemäß Formel 5.29 geschätzt, indem die geschätzte Gesamtanzahl mit dem ermittelten Durchschnittsgewicht g_B einer Beere multipliziert wird. Die Schätzung des Ertrages wird mit dem gewogenen tatsächlichen Ertrag $\mathcal{Y}^T$ der Referenztrauben verglichen. Tabelle 5.8 listet die einzelnen Werte je Referenztraube auf.

	$\mathcal{Y}^T[g]$ pro Traube	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$ pro Traube	$\Delta[\%]$
Traube 13	265,4	285,5	8
Traube 14	86,2	125,4	46
Traube 15	125,2	121,1	-3
Traube 16	217,1	263,8	22
Traube 17	320,0	168,7	-47
Traube 18	261,9	250,9	-4
Traube 19	221,0	207,6	-6
Traube 20	233,7	233,6	0
Traube 21	134,1	142,7	6
Traube 22	222,0	233,6	5
Traube 23	297,1	268,2	-10
Traube 24	302,7	302,8	0
Traube 25	283,3	224,9	-21
$\overline{\text{Traube}}$			-0,3
σ			21
$\sum_{i=13}^{25} \text{Traube}_i$	2969,7	2829,4	-5

Tabelle 5.8: Tatsächlicher Ertrag $\mathcal{Y}^T[g]$ und der geschätzte Ertrag $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$, sowie ihre jeweilige prozentuale Abweichung, die Abweichung ihrer Summen, die mittlere Abweichung und die Standardabweichung für das TS-Erziehungssystem.

Die geschätzten Erträge der einzelnen Referenztrauben erreichen eine Standardabweichung von 21% bei einer mittleren Abweichung von $-0,3\%$. In der Summe beträgt die Abweichung des geschätzten Ertrages vom tatsächlichen Gesamtertrag über alle 13 Referenztrauben -5% .

Mit den Kalibrierwerten f_B und g_B wird eine Schätzung des Gewichtsanteils $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$ aller Beeren der detektierten Trauben mit einer Abweichung von wenigen Prozent vom tatsächlichen Ertrag erreicht. Beide Werte können für die weiteren Berechnungen als repräsentative Kalibrierwerte genutzt werden.

Berechnung der Kalibrierwerte für die MSS-Anbaureihe

Bei der MSS-Anbaureihe ist die Detektion der Beeren aufgrund des vergleichsweise geringeren geometrischen Detailgrades der Rekonstruktion durch Überbelichtung, des höheren Verdeckungsgrades, sowie der Verdeckungen der Kamera selbst durch zu nahe Äste erschwert (s. Kapitel 5.1.2). Nicht alle Referenztrauben können daher für die Berechnung genutzt werden. Vier Referenztrauben wurden nicht detektiert und zwei waren geometrisch falsch rekonstruiert (Abbildung 5.14 zeigt diese zwei falsch rekonstruierten Referenztrauben 2–3). Durch den hohen Verdeckungsgrad sind einige Referenztrauben nur teilweise rekonstruiert worden. Diese Referenztrauben werden jedoch für die Berechnung von f_B genutzt, da sie repräsentativ für den Detektionserfolg der Trauben im MSS-Erziehungssystem sind und somit keine Ausreißer darstellen. Abbildung 5.19 zeigt originale Referenztrauben und die mit dem Berryfinder detektierten Beeren. Eine visuelle Überprüfung zeigt, dass zwar der Großteil der Beeren detektiert wurde und die Kugeln mit tatsächlichen Beeren korrespondieren. Im Vergleich zum TS-Erziehungssystem wurden jedoch weniger vorhandene Beeren detektiert (False Negatives) und mehr Kugeln korrespondieren nicht mit tatsächlichen Beeren (False Positives).

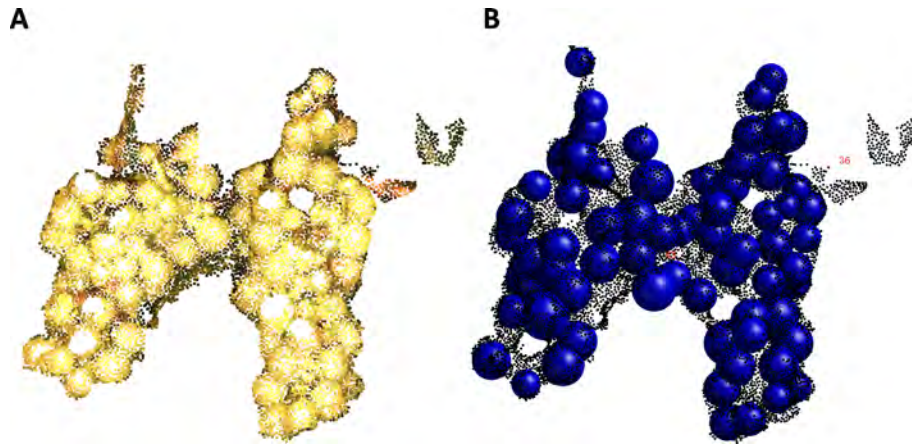


Abbildung 5.19: **A:** Rekonstruierte Referenztrauben der MSS-Anbaureihe. **B:** Mit dem Berryfinder detektierte Beeren in der Referenztraube. Einige Beeren werden nicht detektiert, teilweise korrespondieren Kugeln nicht mit tatsächlichen Beeren.

Die Vorgehensweise bei der Berechnung der Kalibrierwerte ist identisch zur Vorgehensweise bei der TS-Anbaureihe. Sie werden für jede der ersten 12 Referenztrauben berechnet, außer für Referenztrauben, die nicht detektiert wurden oder deren Rekonstruktion fehlerhaft war. Diese sind aufgelistet, die Tabelle enthält für sie jedoch keine Werte, außer zur Berechnung von g_B . Tabelle 5.9 listet die einzelnen Ergebnisse für jede der 12 Referenztrauben auf.

	$\mathcal{D}^B$ pro Traube	$\mathcal{N}^B$ pro Traube	f_B pro Traube	$\mathcal{Y}^T[g]$ pro Traube	$g_B[g]$ pro Beere
Traube 1	20	41	2,05	46,6	1,1
Traube 2	-	104	-	129,8	1,2
Traube 3	-	106	-	133,6	1,3
Traube 4	3	49	16,33	50,7	1,0
Traube 5	14	44	3,93	66	1,2
Traube 6	21	32	1,52	22,9	0,7
Traube 7	21	51	2,43	47,3	0,9
Traube 8	-	41	-	48,9	1,2
Traube 9	16	61	3,81	45,6	0,7
Traube 10	11	31	2,82	32,3	1,0
Traube 11	22	53	2,41	93,3	1,7
Traube 12	29	56	1,93	56,4	1,0
Traube	$m_B = 20$	$B_T = 51$	$f_B = 2,43$	$g_B = 1,1$	

Tabelle 5.9: Berechnung des Kalibrierwertes f_B , sowie des Medians der detektierten Beerenanzahl m_B , der tatsächlichen Beerenanzahl B_T und des durchschnittlichen Gewichts einer Beere g_B aus den Referenzwerten der MSS-Anbaureihe. Werte sind gerundet dargestellt.

Der Median der durchschnittlich detektierten, sichtbaren Beeren ergibt sich zu $m_B = 20$. Die durchschnittliche tatsächliche Anzahl der Beeren an einer Traube ergibt sich zu $B_T = 51$. Das Gewicht einer Einzelbeere beträgt $g_B = 1,1g^4$. Es ergibt sich ein Kalibrierwert $f_B = 2,43$, der für die weiteren Berechnungen in den MSS-Anbaureihen als Kalibrierwert verwendet wird, wobei die Notation $f_B = \widehat{f}_B$ gilt.

Für die verbleibenden 10 der ursprünglich 13 unabhängigen Referenztrauben wird $\hat{\mathcal{E}}_V^B$ pro Traube berechnet. Diese werden in Tabelle 5.10 mit den Referenzwerten $\mathcal{N}^B$ pro Traube der 10 Referenztrauben verglichen.

⁴Der Wert wird auf ebenfalls leicht auf $1[g]$ reduziert, um das Gewicht des Stielgerüsts zu kompensieren.

	$\mathcal{N}^B$ pro Traube	$\hat{\mathcal{E}}_V^B$ pro Traube	$\Delta[\%]$
Traube 13	55	90	63
Traube 14	-	-	-
Traube 15	-	-	-
Traube 16	48	61	26
Traube 17	83	39	-53
Traube 18	48	49	1
Traube 19	49	24	-50
Traube 20	-	-	-
Traube 21	61	58	-4
Traube 22	81	53	-34
Traube 23	99	83	-17
Traube 24	26	22	-16
Traube 25	122	49	-60
$\overline{\text{Traube}}$			14
σ			39
$\sum_{i=13}^{25} \text{Traube}_i$	672	528	-22

Tabelle 5.10: Tatsächliche Beerenanzahl $\mathcal{N}^B$, geschätzte Beerenanzahl $\hat{\mathcal{E}}_V^B$ pro Traube, ihre prozentuale Abweichung, die Abweichung ihrer Summen, die mittlere Abweichung und die Standardabweichung für das MSS-Erziehungssystem.

Die Standardabweichung einer Schätzung der Beerenanzahl beträgt 39% für das MSS-Erziehungssystem, bei einer mittleren Abweichung von 14%. Die Summe von $\hat{\mathcal{E}}_V^B$ der 10 Referenztrauben wird um 22% geringer als die Summe von $\mathcal{N}^B$ geschätzt.

	$\mathcal{Y}^T[g]$ pro Traube	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$ pro Traube	$\Delta[\%]$
Traube 13	69,9	90,0	29
Traube 14	-	-	-
Traube 15	-	-	-
Traube 16	45,9	61,0	33
Traube 17	114,1	39,0	-66
Traube 18	47,1	49,0	4
Traube 19	59,1	24,0	-59
Traube 20	-	-	-
Traube 21	64,4	58,0	-10
Traube 22	90,8	53,0	-42
Traube 23	139,8	83,0	-41
Traube 24	21,5	22,0	2
Traube 25	136,8	49	-64
$\overline{\text{Traube}}$			-21
σ			38
$\sum_{i=13}^{25} \text{Traube}_i$	789,4	528,0	-33

Tabelle 5.11: Tatsächlicher Ertrag $\mathcal{Y}^T[g]$, geschätzter Ertrag $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$ pro Referenztraube, ihre prozentuale Abweichung, die Abweichung ihrer Summen, mittlere Abweichung und Standardabweichung für das MSS-Erziehungssystem.

Der Ertrag der 10 detektierten Referenztrauben $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ wird nach Formel 5.29 geschätzt und mit dem Ertrag $\mathcal{Y}^T$ der Referenztrauben verglichen. Tabelle 5.11 listet die einzelnen Werte je Referenztraube, ihre prozentuale Abweichung, die Abweichung ihrer Summen und die Standardabweichung auf.

Die Schätzung der Erträge der einzelnen Referenztrauben weisen eine Standardabweichung von 38% bei einer mittleren Abweichung von -21% auf. In der Summe beträgt die Abweichung vom tatsächlichen Ertrag über alle 10 detektierten Referenztrauben -33% . Mit den Kalibrierwerten f_B und g_B ist eine Schätzung des Gewichtsanteils $\hat{y}(\hat{\mathcal{E}}_V^B)$ der sichtbaren Beeren zwar möglich, jedoch mit größerer Abweichung vom tatsächlichen Ertrag als bei den TS-Anbaureihen. Dies resultiert aus dem hohen Verdeckungsgrad im MSS-Erziehungssystem und den daher teilweise unvollständigen Punktwolken der Trauben. Dadurch können 1. weniger Beeren detektiert werden und 2. ist der Grad der Vollständigkeit der Trauben unbeständig. Durch diese Unbeständigkeit ist f_B nur bedingt repräsentativ für individuelle Trauben. $\hat{\mathcal{E}}_V^B$ besitzt daher eine größere Standardabweichung bzgl. der Einzelmessungen als im Fall der TS-Anbaureihe und die Abweichung vom Ertrag ist entsprechend höher.

5.3.2.2 Kalibrierwerte für die Anzahl nicht detektierter Trauben

Im Folgenden werden die Kalibrierwerte $\hat{\mathcal{M}}_u^T$ und f_m zur Schätzung der nicht detektierten Trauben $\mathcal{M}^T$ beschrieben und ihre Berechnung vorgestellt.

Nicht detektierte Trauben zählen wie in Kapitel 4.3.2 erläutert als False Negatives. Dort wurden fünf Gründe angeführt, warum Trauben nicht detektiert werden. Diese Gründe lassen sich in zwei Kategorien einteilen, die zur Berechnung von $\hat{\mathcal{M}}_u^T$ und f_m erläutert werden müssen:

1. Anzahl der False Negatives, die aus der Vereinigung mehrerer Trauben in einer Komponente resultieren: Diese Anzahl sei im Weiteren als $\mathcal{M}_u^T$ (united) bezeichnet. Die Vereinigung der Trauben verfälscht zwar die Zählung der Trauben, jedoch wurden die sie enthaltende Komponente als Traube akzeptiert und *nicht* von der Auswertemethode entfernt. Die Komponente kann daher mit dem Berryfinder nach Beeren durchsucht werden. Somit ist basierend auf konkreten Zählungen der Beeren in dieser Komponente abschätzbar, wie groß die Anzahl der in dieser Komponente vereinigten Trauben ist. $\hat{\mathcal{M}}_u^T$ ist die Approximation von $\mathcal{M}_u^T$.
2. Anzahl $\mathcal{M}_l^T$ (lost) der für den Sensor sichtbaren, aber von der Auswertemethode entfernten Trauben. Diese Anzahl umfasst all jene Trauben, die entweder nicht rekonstruiert oder nicht klassifiziert wurden oder jene Trauben die verworfen wurden, weil ihre Punktzahl weniger als p_K betrug oder weil in ihnen weniger als $\min_B$ Beeren gefunden wurden. Diese Trauben wurden *nicht als Traube erkannt* und folglich aus den weiter zu untersuchenden Daten entfernt. Die Detektion von Beeren ist bei ihnen daher *nicht* möglich. Die Anzahl dieser Trauben muss daher über den Kalibrierwert f_m approximiert werden.

$\mathcal{M}^T$ ist die Summe dieser Teilanzahlen (Formel 5.15). Für die ersten drei Abschnitte wird die Zusammensetzung von $\mathcal{M}^T$ durch manuelle Überprüfung in den Bildern und Punktwolken nach Gründen für die Nicht-Detektion pro Abschnitt aufgeschlüsselt (vgl. Tabelle 4.12 in Kapitel 4.3.2). Dadurch werden Referenzwerte $\mathcal{M}_u^T$ und $\mathcal{M}_l^T$ erhalten. Anhand dieser wird überprüft, ob $\mathcal{M}_u^T$ über $\hat{\mathcal{M}}_u^T$ korrekt approximiert werden kann. Des Weiteren wird f_m für die TS- und MSS-Anbaureihen berechnet. Die Vorgehensweise wird im Folgenden geschildert.

Abschätzung der Anzahl nicht trennbarer Trauben $\hat{\mathcal{M}}_u^T$

Komponenten, in denen mehr als 50 Beeren detektiert wurden, vereinen höchstwahrscheinlich mehrere Trauben in sich. Dies ergibt sich aus den ermittelten Durchschnittswerten der detektierten Beerenanzahl pro Traube m_B für jedes Erziehungssystem. Zur Bestimmung von $\hat{\mathcal{M}}_u^T$ wird $\mathcal{D}^B$ in den zu prüfenden Komponenten durch m_B dividiert (Formel 5.16). $\hat{\mathcal{M}}_u^T$ wird zu $\mathcal{D}^T$ addiert und die darum korrigierte Anzahl $\hat{\mathcal{D}}_c^T$ (corrected) der sichtbaren Trauben gemäß Formel 5.17 erhalten.

Abschätzung von $\hat{\mathcal{M}}_l^T$ über f_m

Mit f_m wird anschließend aus $\hat{\mathcal{D}}_c^T$ die Anzahl $\hat{\mathcal{M}}_l^T$ der entfernten Trauben geschätzt. Die Berücksichtigung von $\hat{\mathcal{M}}_l^T$ über einen Kalibrierwert kann aus den Ergebnissen der Voruntersuchungen (Kapitel 4.3.2) als adäquates Mittel angesehen werden. In diesem Kapitel konnte gezeigt werden, dass der durchschnittliche Recall für das BBCH89 Reifestadium bei 71% und 68% für das TS- bzw. das MSS-Erziehungssystem lag. Der

Anteil nicht detektierter Trauben $\mathcal{M}^T$, unabhängig von Reifestadium und Erziehungssystem, lag über fünf Meter bei ungefähr einem Drittel, was mit dem durchschnittlichen Recall korrespondiert. Die Ursachen für die Nicht-Detektion der Trauben nahmen dabei Werte an, die jeweils typische Muster für Reifestadium und Erziehungssystem darstellen. Die Verwendung eines Kalibrierwertes erscheint daher plausibel, um die nicht detektierten Trauben zu berücksichtigen. Der Kalibrierwert f_m ist das Verhältnis der manuell gezählten, tatsächlichen Anzahl der nicht detektierten Trauben $\mathcal{M}_l^T$ und $\hat{\mathcal{D}}_c^T$ gemäß Formel 5.18, Tabelle 5.3. Die Schätzung von $\hat{\mathcal{M}}_l^T$ wird durch Multiplikation von f_m mit $\hat{\mathcal{D}}_c^T$ nach Formel 5.19, Tabelle 5.3 berechnet.

Überprüfung von $\hat{\mathcal{M}}_u^T$ und Berechnung von f_m für die TS-Anbaureihe

Zur Berechnung von f_m wird zunächst die Anzahl $\mathcal{D}^T$ bestimmt. Abbildung 5.20 illustriert detektierte Trauben des Abschnitts 3. Tabelle 5.12 gibt einen Überblick über $\mathcal{M}^T$:

	(I) $\mathcal{D}^T$	(II) $\mathcal{M}^T$	(1) Nicht rekonstruiert	(2) Falsch klassifiziert	(3) Punktzahl $< p_K$	(4) Beerenzahl $< min_B$	(5) Nicht trennbar
Meter 1	21	13	1	4	0	2	6
Meter 2	12	17	0	0	0	1	16
Meter 3	16	21	0	1	0	1	19

Tabelle 5.12: $\mathcal{D}^T$ und $\mathcal{M}^T$ in den Spalten (I) und (II), sowie die Zusammensetzung von $\mathcal{M}^T$ in absoluten Zahlen für die Abschnitte 1 – 3 der TS-Anbaureihe.

Die Detektion von Traubenkandidaten mittels des CC-Verfahrens erweist sich auch im Falle der Phenoliner Daten als adäquate Strategie zur Detektion von Trauben, erzeugt jedoch im Vergleich zu Daten des PHE-NObots deutlich mehr Komponenten $\mathcal{M}_u^T$, in denen mehrere Trauben vereint sind (vgl. Tabelle 4.12 und Tabelle 5.12).

$\mathcal{M}_u^T$ (Spalte 5) ist auch deutlich höher, als die Anzahl $\mathcal{M}_l^T$ (Tabelle 5.12, Spalten 1-4). An den Rändern der Trauben liegen aufgrund von Mischpixeln aus Vorder- und Hintergrund (s. Kapitel 5.1.4) falsch rekonstruierte Punkte. Durch die Mischpixel sind diese dunkler als der übrige Hintergrund und konnten während der Vorverarbeitung der Punktwolken nicht entfernt werden. Dadurch existiert nur ein geringer oder gar kein Zwischenraum zwischen benachbarten Trauben (s. Abbildungen 5.11 und 5.20). Da Trauben und Mischpixel teilweise ähnlich gefärbt sind, werden diese auch im weiteren Verlauf der Auswertung fälschlicherweise als Traube klassifiziert. Die räumliche Trennung der Trauben mittels des CC-Verfahrens wird dadurch erschwert.

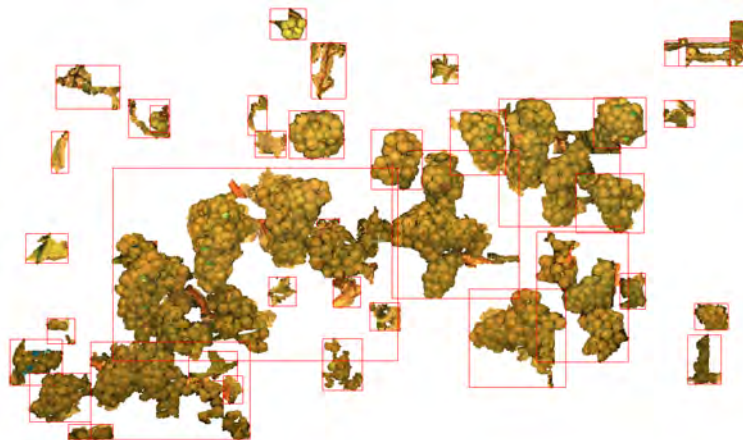


Abbildung 5.20: Ausschnitt der Punktwolke aus Abschnitt 3 mit Traubenkandidaten. Komponenten sind durch eine rote Bounding Box gekennzeichnet. Es können mehrere Trauben erkannt werden, die in einer Komponente vereint sind. Bei einigen Trauben fehlen kleinere Teile. Des Weiteren sind einige Traubenkandidaten vorhanden, die keiner tatsächlichen Traube entsprechen, sondern bspw. Blattreste oder Teile der Verdrahtung darstellen.

Die Anzahl $\hat{\mathcal{M}}_l^T$ ist im Vergleich mit den PHENObot Daten hingegen geringer (s. Kapitel 4, Tabelle 4.12). Es werden deutlich weniger tatsächliche Trauben aufgrund einer Punktzahl $< p_K$ oder einer Beerenanzahl $< \min_B$ entfernt. Diese Verbesserung ist vermutlich auf den geringeren Verdeckungsgrad des ausgedünnten Blattwerks, der eine vollständigere Rekonstruktion der Trauben begünstigt, sowie die erhöhte Bildabdeckung durch eine kurze $base_h$ zurückzuführen.

Die Anzahl $\hat{\mathcal{M}}_u^T$ wird nach Formel 5.16 für jeden der drei Abschnitte geschätzt. Als Median der detektierten Beerenanzahl wird der Wert aus Tabelle 5.6 mit $m_B = 43$ genutzt.

Die Gültigkeit des Vorgehens wird überprüft, indem zunächst diejenigen Komponenten manuell bestimmt werden, in denen tatsächlich mehrere Trauben vereint sind. Anschließend wird geprüft, ob diese Komponenten durch das Vorgehen identifiziert werden konnten und ob $\hat{\mathcal{M}}_u^T$ der tatsächlichen Anzahl $\mathcal{M}_u^T$ entspricht (Tabelle 5.12, letzte Spalte). $\mathcal{D}^T$ wird anschließend durch Addition von $\hat{\mathcal{M}}_u^T$ korrigiert und $\hat{\mathcal{D}}_c^T$ erhalten. Tabelle 5.13 listet die Ergebnisse auf:

	Komp.	$\mathcal{M}_u^T$	$\hat{\mathcal{M}}_u^T$	$\hat{\mathcal{D}}_c^T$
Meter 1	3	6	6	27
Meter 2	5	16	16	28
Meter 3	5	19	15	31

Tabelle 5.13: Anzahl der Komponenten mit mehr als 50 Beeren, die Anzahl der in diesen Komponenten enthaltenen tatsächlichen Traubenanzahl $\mathcal{M}_u^T$, die Approximation dieser Anzahl $\hat{\mathcal{M}}_u^T$ und die um $\hat{\mathcal{M}}_u^T$ korrigierte Traubenanzahl $\mathcal{D}_c$.

Alle Komponenten, die mehrere Trauben enthielten, konnten identifiziert werden. In keiner überprüften Komponente, die richtigerweise nur eine Traube enthielt, wurde fälschlicherweise mehr als eine Traube geschätzt. Die Ergebnisse demonstrieren, dass die tatsächliche Anzahl der in einer Komponente vereinten Trauben zum Großteil korrekt abgeschätzt wird. Lediglich in Abschnitt 3 wich die Zählung um vier Trauben von der tatsächlichen Anzahl ab.

Anhand der korrigierten Anzahl $\hat{\mathcal{D}}_c^T$ aus Tabelle 5.13 wird der Kalibrierwert f_m (Formel 5.18) berechnet. Für jeden der drei Abschnitte wird f_m und daraus der Mittelwert bestimmt. Tabelle 5.14 listet die Ergebnisse auf.

	$\hat{\mathcal{D}}_c^T$	$\mathcal{M}_l^T$	f_m
Meter 1	27	7	0,3
Meter 2	28	1	0,03
Meter 3	31	2	0,06
$\overline{f_m}$			0,1

Tabelle 5.14: Geschätzte Anzahl $\hat{\mathcal{D}}_c^T$, gezählte tatsächliche Anzahl $\mathcal{M}_l^T$ der False Negatives und der jeweilige Kalibrierwert f_m pro Abschnitt, sowie sein Mittelwert aus den drei Abschnitten des TS-Erziehungssystems.

Der mittlere Kalibrierwert $\overline{f_m}$ wird als Kalibrierwert für die weitere Berechnung der TS-Anbaureihe verwendet, wobei die Notation $f_m = \overline{f_m}$ genutzt wird.

Überprüfung von $\hat{\mathcal{M}}_u^T$ und Berechnung von f_m für die MSS-Anbaureihe

Zur Berechnung des Kalibrierwertes f_m für die MSS-Anbaureihe wird identisch wie für die TS-Anbaureihe vorgegangen. Tabelle 5.15 listet die Zusammensetzung von $\mathcal{M}^T$ für die Abschnitte des MSS-Erziehungssystems auf. Wie schon in den Untersuchungen mit dem PHENObot werden bei der MSS-Anbaureihe Trauben vermehrt aufgrund des hohen Verdeckungsgrades nicht rekonstruiert. Keine vollständige Traube wurde falsch klassifiziert, lediglich Teile der Trauben wurden stellenweise falsch klassifiziert.

Deutlich mehr Trauben im MSS-Erziehungssystem erreichen jedoch im Vergleich zum TS-Erziehungssystem nicht die nötige Mindestmenge an Beeren $\min_B$. Dies hat zwei Gründe. Aufgrund der höheren Verdeckung

durch Blätter können Trauben nur teilweise rekonstruiert werden oder Teile der Traube wurden falsch klassifiziert, so dass weniger Beeren überhaupt detektierbar sind. Die zweite Ursache ist die geringere geometrische Genauigkeit der Rekonstruktion. In einer teilweise rekonstruierten Traube können bspw. ausreichend Beeren vorhanden sein, die jedoch aufgrund einer korruptierten Geometrie nicht vom Berryfinder detektiert werden können.

	(I)	(II)	(1)	(2)	(3)	(4)	(5)
	$\mathcal{D}^T$	$\mathcal{M}^T$	Nicht rekonstruiert	Falsch klassifiziert	Punktzahl $< p_K$	Beerenanzahl $< min_B$	Nicht trennbar
Meter 1	9	5	2	0	1	0	2
Meter 2	7	5	2	0	0	3	0
Meter 3	11	10	3	0	0	5	2

Tabelle 5.15: $\mathcal{D}^T$ und $\mathcal{M}^T$ in den Spalten (I) und (II), sowie die ursächliche Zusammensetzung von $\mathcal{M}^T$ in absoluten Zahlen für die Abschnitte 1 – 3 der MSS-Anbaureihe.

In MSS-Erziehungssystem werden nur wenige Trauben durch das CC-Verfahren in einer einzigen Komponente zusammengefasst. Diese Anzahl ist im Vergleich zum TS-Erziehungssystem bedeutend geringer, da im MSS-Erziehungssystem Trauben in großem räumlichen Abstand zueinander wachsen (s. bspw. Abbildung 4.35). Die Anzahl $\hat{\mathcal{M}}_u^T$ wird nach Formel 5.16 für jeden der drei Abschnitte berechnet. Als Median der detektierten Beerenanzahl wird der Wert aus Tabelle 5.9 mit $m_B = 20$ verwendet. Tabelle 5.16 listet die Ergebnisse auf:

	Komp.	$\mathcal{M}_u^T$	$\hat{\mathcal{M}}_u^T$	$\hat{\mathcal{D}}_c^T$
Meter 1	1	2	4	13
Meter 2	0	0	0	7
Meter 3	1	2	3	14

Tabelle 5.16: Anzahl der Komponenten mit mehr als 50 Beeren, die Summe der in diesen Komponenten enthaltenen tatsächlichen Traubenanzahl $\mathcal{M}_u^T$, die Approximation dieser Anzahl $\hat{\mathcal{M}}_u^T$ und die um $\hat{\mathcal{M}}_u^T$ korrigierte Traubenanzahl $\hat{\mathcal{D}}_c$ für das MSS-Erziehungssystem.

Alle Komponenten, die mehrere Trauben enthielten, konnten identifiziert werden. In keiner überprüften Komponente, die richtigerweise nur eine Traube enthielt, wurde fälschlicherweise mehr als eine Traube geschätzt. Die Anzahl der Trauben wird in Abschnitt 1 um zwei Trauben, in Abschnitt 3 um eine Traube überschätzt. Dies resultiert aus den oftmals nur teilweise vorhandenen Trauben, wodurch der Durchschnitt detektierter Trauben m_B kleiner wird. Sind Trauben vollständiger als die meisten Trauben vorhanden und werden in einer Komponenten zusammengefasst, ist m_B zu klein, um die korrekte Anzahl der Trauben abzuschätzen. Stattdessen wird die Anzahl überschätzt. In Abschnitt 2 wurden korrekterweise keine zusammenhängenden Trauben identifiziert.

	$\hat{\mathcal{D}}_c^T$	$\mathcal{M}_l^T$	f_m
Meter 1	13	3	0,2
Meter 2	7	5	0,7
Meter 3	14	8	0,5
$\overline{f_m}$			0,5

Tabelle 5.17: Geschätzte Anzahl sichtbarer Trauben $\hat{\mathcal{D}}_c^T$, gezählte tatsächliche Anzahl $\mathcal{M}_l^T$ der False Negatives und der jeweilige Kalibrierwert f_m , sowie sein Mittelwert aus den drei Abschnitten für das MSS-Erziehungssystem.

Anhand der korrigierten Anzahl $\hat{\mathcal{D}}_c^T$ wird der Kalibrierwert f_m (Formel 5.18) für jeden der ersten drei Abschnitte bestimmt und der Mittelwert berechnet. Tabelle 5.17 listet die Ergebnisse auf. Der mittlere Kalibrierwert $\overline{f_m} = 0,5$ wird als Kalibrierwert für die weitere Berechnung verwendet, wobei die Notation $f_m = \overline{f_m}$ genutzt wird.

5.3.2.3 Kalibrierwert für die Anzahl stets verdeckter Trauben

Der Kalibrierwert f_o wird entsprechend Formel 5.21, Tabelle 5.3 berechnet. Dazu wird zunächst die geschätzte Anzahl der sichtbaren Trauben $\hat{\mathcal{V}}^T$ nach Formel 5.20 bestimmt. Der Kalibrierwert ist das Verhältnis von $\hat{\mathcal{V}}^T$ und der tatsächlichen Anzahl nicht sichtbarer Trauben $\mathcal{O}^T$. Die geschätzte Anzahl $\hat{\mathcal{O}}^T$ ergibt sich durch Multiplikation von $\hat{\mathcal{V}}^T$ mit f_o nach Formel 5.22, Tabelle 5.3.

Berechnung von f_o für die TS-Anbaureihe

In den drei Abschnitten der TS-Anbaureihe konnten keine Trauben festgestellt werden, die von jeder Seniorsperspektive aus verdeckt waren. Alle Trauben war stets mindestens anteilig sichtbar. Der Kalibrierwert f_o ergibt sich daher zu $f_o = 0$. Es ist jedoch zu erwarten, dass diese Zahl für TS-Anbaureihen ohne vorherige Ausdünnung des Blattwerks höher liegen wird.

Berechnung von f_o für die MSS-Anbaureihe

Der Kalibrierwert f_o wird für jeden der drei Abschnitte berechnet und der Mittelwert berechnet. Tabelle 5.18 listet die Ergebnisse auf:

	$\hat{\mathcal{V}}^T$	$\mathcal{O}^T$	f_o
Meter 1	20	51	2,6
Meter 2	11	37	3,5
Meter 3	21	38	1,8
$\overline{f_o}$			2,6

Tabelle 5.18: Geschätzte Anzahl sichtbarer Trauben $\hat{\mathcal{V}}^T$, tatsächliche Anzahl $\mathcal{O}^T$ stets verdeckter Trauben und der jeweilige Kalibrierwert f_o , sowie sein Mittelwert aus den drei Abschnitten für das MSS-Erziehungssystem.

Der mittlere Kalibrierwert $\overline{f_o} = 2,6$ wird im Folgenden zur Schätzung der Anzahl stets verdeckter Trauben für die MSS-Anbaureihe verwendet, wobei die Notation $f_o = \overline{f_o}$ verwendet wird.

Bedeutung von f_o für das MSS-Erziehungssystem

Das Verhältnis von $\mathcal{O}^T$ und $\hat{\mathcal{V}}^T$ ist für das MSS-Erziehungssystem von besonderer Bedeutung für die Ertragsschätzung. $\mathcal{O}^T$ ist stets deutlich größer als $\hat{\mathcal{V}}^T$, ihr Verhältnis beträgt teilweise 1:3. Dies bedeutet, dass der Erfolg der Ertragsschätzung primär von der Schätzung des *unsichtbaren* und nicht von der Detektion des sichtbaren Teils der Anbaureihe abhängt.

Die Sicherheit der Schätzung von $\hat{\mathcal{O}}^T$ über f_o hängt neben der Beständigkeit von Recall und Precision der Auswertemethode vor allem von der Gleichförmigkeit der Weinreben und der Anzahl der an ihr wachsenden Trauben ab. Nur wenn die Anzahl unsichtbarer Trauben stets weitestgehend proportional zur Anzahl der detektierten Trauben ist, kann ein für alle Abschnitte repräsentativer Kalibrierwert ermittelt werden. Die Verwendung eines Kalibrierwertes führt aufgrund der Variabilität von Weinreben daher eine große Unsicherheit in die Bestimmung von $\hat{\mathcal{O}}^T$ ein. Ob der sichtbare Teil der Anbaureihe korrekt und vollständig detektiert wird, spielt eine vergleichsweise geringere Rolle, da sein Anteil am Gesamtgewicht geringer ist.

5.3.3 Schätzung der Gesamtanzahl aller Beeren und Trauben

Die Summe der Teilanzahlen $\hat{\mathcal{D}}_c^T$, $\hat{\mathcal{M}}_l^T$ und $\hat{\mathcal{O}}^T$ ist die geschätzte Gesamtanzahl aller Trauben $\hat{\mathcal{E}}^T$ (Formel 5.23, vgl. Abbildung 5.17). Mit der geschätzten Anzahl aller Beeren der sichtbaren Trauben $\hat{\mathcal{E}}_V^B$ und der geschätzten Anzahl aller nicht detektierten und der stets unsichtbaren Trauben $\hat{\mathcal{M}}_l^T$ bzw. $\hat{\mathcal{O}}^T$ lässt sich die Gesamtanzahl der Beeren berechnen. Dabei wird auf die mittlere Beerenanzahl pro Traube B_T zurückgegriffen, die in Kapitel 5.3.2.1 berechnet wurde. Man erhält die Beerenanzahl $\hat{\mathcal{M}}_l^B$ und $\hat{\mathcal{O}}^B$ der Trauben $\hat{\mathcal{M}}_l^T$ und $\hat{\mathcal{O}}^T$, indem sie mit B_T multipliziert werden (Formeln 5.26 und 5.27). Die Gesamtanzahl aller Beeren $\hat{\mathcal{E}}^B$ ist die Summe der Teilanzahlen $\hat{\mathcal{E}}_V^B$, $\hat{\mathcal{M}}_l^B$ und $\hat{\mathcal{O}}^B$ (Formel 5.28).

5.3.4 Funktionsvorschrift zur Schätzung des Ertrages

Die zu Beginn des Kapitels skizzierte Funktion 5.14 zur Schätzung des Ertrages ist eine Funktion der Beeren und Trauben:

- Die Anzahl $\mathcal{D}^B$ stellt eine gute Zählung des sichtbaren Teils der Beeren aller sichtbaren Trauben dar (s. Kapitel 4.3.2). Die geschätzte Anzahl $\hat{\mathcal{E}}_V^B$ gilt dabei als Anzahl aller Beeren des detektierten, sichtbaren Traubenbestandes und berücksichtigt damit implizit die Anzahl der sichtbaren Trauben. Einzelne Beeren stellen den kleinsten zählbaren Gewichtsanteil einer Traube dar. Eine konkrete Zählung der Beeren multipliziert mit dem durchschnittlichen Gewicht pro Beere g_B ermöglicht daher eine präzisere Gewichtsabschätzung als die geschätzte Anzahl der Trauben $\hat{\mathcal{D}}_c^T$ multipliziert mit einem durchschnittlichen Traubengewicht. Grundsätzlich sind daher Gewichtsabschätzungen über die Beeren zu bevorzugen, wenn für letztere konkrete Zählungen vorliegen. Die geschätzte Anzahl der sichtbaren Trauben $\hat{\mathcal{D}}_c^T$ wird daher explizit *nicht* zur Schätzung des Ertrages der sichtbaren und detektierten Trauben verwendet. Stattdessen wird ihr Gewicht über die Anzahl ihrer Beeren $\hat{\mathcal{E}}_V^B$ geschätzt.
- Um eine realistische Ertragsschätzung möglichst nah am tatsächlichen Ertrag zu erhalten, müssen die Anzahlen $\hat{\mathcal{M}}_l^T$ und $\hat{\mathcal{O}}^T$ ebenfalls bei der Ertragsschätzung berücksichtigt werden. Sie werden basierend auf $\mathcal{D}^T$ abgeschätzt und haben damit ebenfalls einen Bezug zu konkreten Zählungen. Ihr Ertrag wird über ihre geschätzten Beerenanzahlen $\hat{\mathcal{M}}_l^B$ und $\hat{\mathcal{O}}^B$ bestimmt, indem sie mit dem durchschnittlichen Gewicht einer Einzelbeere g_B multipliziert werden.

Der geschätzte Ertrag setzt sich diesen Ausführungen folgend aus dem Ertrag der Teilanzahlen $\hat{\mathcal{E}}_V^B$ der Beeren der sichtbaren Trauben sowie dem Ertrag der Beeren der nicht detektierten Trauben $\hat{\mathcal{M}}_l^B$ und dem Ertrag der Beeren der stets verdeckten Trauben $\hat{\mathcal{O}}^B$ zusammen (Formeln 5.29, 5.30, 5.31 in Tabelle 5.5, vgl. Abbildung 5.17). Die Funktionsvorschriften zur Berechnung der Gewichtsanteile der Teilanzahlen lauten diesen Formel folgend vollständig:

$$\begin{aligned}\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B) &= \mathcal{D}^B \cdot f_B \cdot g_B \\ &= \hat{\mathcal{E}}_V^B \cdot g_B\end{aligned}\tag{5.33}$$

$$\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B) = \hat{\mathcal{D}}_c^T \cdot f_m \cdot B_T \cdot g_B\tag{5.34}$$

$$\begin{aligned}&= \hat{\mathcal{M}}_l^T \cdot B_T \cdot g_B \\ &= \hat{\mathcal{M}}_l^B \cdot g_B\end{aligned}\tag{5.35}$$

$$\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B) = \hat{\mathcal{V}}^T \cdot f_o \cdot B_T \cdot g_B\tag{5.36}$$

$$\begin{aligned}&= \hat{\mathcal{O}}^T \cdot B_T \cdot g_B \\ &= \hat{\mathcal{O}}^B \cdot g_B.\end{aligned}\tag{5.37}$$

Der Gesamtertrag ist die Summe der Teilerträge der Teilanzahlen nach Formel 5.32. Da die Teilerträge über die Anzahl der Beeren bestimmt werden, entspricht der Ertrag dem Gewicht aller Beeren (vgl. Abbildung 5.17):

$$\hat{\mathcal{Y}} = \hat{\mathcal{E}}^B \cdot g_B\tag{5.38}$$

5.3.5 Schätzung des Ertrages der fünf Abschnitte beider Erziehungssysteme

Alle notwendigen Größen für die Ertragsschätzung beider Erziehungssysteme wurden berechnet. Die Ertragsparameter und der Ertrag werden im folgenden Unterkapitel für alle fünf Abschnitte nach der oben

beschriebenen Vorgehensweise geschätzt⁵. Zunächst werden die Teilanzahlen $\hat{\mathcal{E}}_V^B$, $\hat{\mathcal{D}}_c^T$, $\hat{\mathcal{M}}_l^T$ und $\hat{\mathcal{O}}^T$ aus $\mathcal{D}^B$ und $\mathcal{D}^T$ pro Abschnitt bestimmt. Um den Erfolg der Ertragsschätzung zu überprüfen, wird der Ertrag $\hat{\mathcal{Y}}$ pro Abschnitt geschätzt und mit dem tatsächlichen Ertrag $\mathcal{Y}^A$ pro Abschnitt und in der Summe verglichen. Dabei ist zu beachten, dass die Ergebnisse für die ersten drei Abschnitte mit den bestimmten Kalibrierwerten korreliert sind. Die Abschnitte vier und fünf hingegen stellen von den Kalibrierwerten völlig unabhängige Referenzwerte dar.

Die Ertragsschätzung wird anhand von Weinreben mit nur wenig verdeckten Trauben (TS-Anbaureihe) und anhand von Weinreben mit einem hohen Verdeckungsgrad (MSS-Anbaureihe) evaluiert. Das Ausmaß des Verdeckungsgrades ist ein *elementarer* Einflussfaktor bei der Ertragsschätzung. Verdeckungen erschweren die Beobachtung der Trauben oder verhindern sie gänzlich. Verdeckungen lassen sich, unabhängig vom Erziehungssystem, in der weinbaulichen Ertragsschätzung nie vollständig vermeiden. Je mehr Trauben dauerhaft außerhalb des Sichtbereichs der Kameras liegen, desto unsicherer wird die Ertragsschätzung. Es ist daher wichtig, Erkenntnisse über den Einfluss der Verdeckungen auf die Ertragsschätzung zu sammeln und inwieweit der Einfluss der Verdeckung durch Kalibrierwerte berücksichtigt werden kann. Zu erwarten ist, dass die die Ertragsschätzung für die MSS-Anbaureihe aufgrund der großen Anzahl unsichtbarer Trauben unsicherer sein wird, als für die TS-Anbaureihe.

Tabelle 5.19 listet die Teilanzahlen der Beeren und Trauben auf, die für die Schätzung des Ertrages nach den Formeln 5.29 - 5.31 benötigt werden. Abbildung 5.21 zeigt beispielhaft das Ergebnis der Beerendetektion für einen Ausschnitt des Abschnitts 3 der TS-Anbaureihe. Abbildung 5.22 zeigt mehrere Trauben und darin detektierte Beeren aus Abschnitt 3 der MSS-Anbaureihe.

	TS				MSS			
	$\hat{\mathcal{E}}_V^B$	$\hat{\mathcal{D}}_c^T$	$\hat{\mathcal{M}}_l^T$	$\hat{\mathcal{O}}^T$	$\hat{\mathcal{E}}_V^B$	$\hat{\mathcal{D}}_c^T$	$\hat{\mathcal{M}}_l^T$	$\hat{\mathcal{O}}^T$
Meter 1	2320	27	3	0	426	13	7	52
Meter 2	2849	28	3	0	264	7	4	28
Meter 3	2777	31	4	0	530	14	7	56
Meter 4	2203	21	3	0	271	9	5	36
Meter 5	2030	24	3	0	336	8	4	32

Tabelle 5.19: Geschätzte Anzahl $\hat{\mathcal{E}}_V^B$ aller Beeren aller sichtbaren und detektierten Trauben, sowie die geschätzten Traubenanzahlen der sichtbaren Trauben $\hat{\mathcal{D}}_c^T$, der sichtbaren, aber nicht detektierten $\hat{\mathcal{M}}_l^T$ und der stets unsichtbaren Trauben $\hat{\mathcal{O}}^T$ pro Abschnitt. Die Linie zwischen den Abschnitten 1 – 3 und 4 – 5 trennt korrelierte Ergebnisse von unabhängigen Ergebnissen.

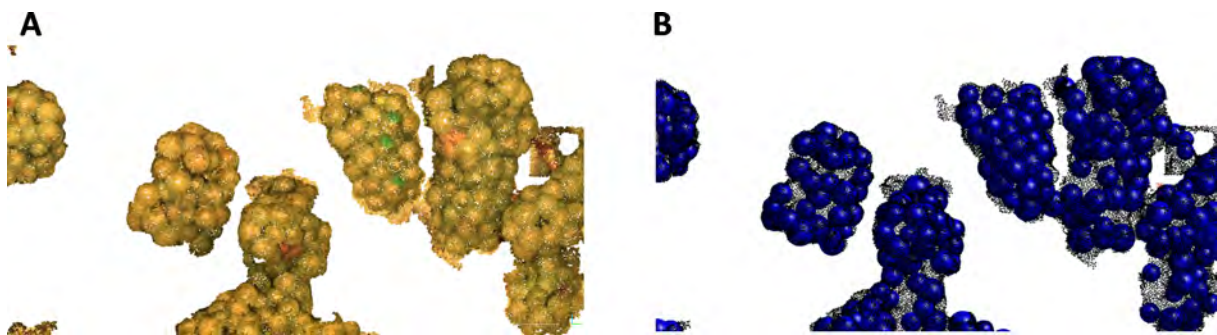


Abbildung 5.21: **A:** Ausschnitt der Punktwolke aus Abschnitt 3 der TS-Anbaureihe. **B:** Mit dem Berryfinder detektierte Beeren im dargestellten Ausschnitt.

⁵Es wird darauf verzichtet, erneut wie in Kapitel 4.3, Recall und Precision von $\mathcal{D}^T$ und $\mathcal{D}^B$ für die Phenoliner Daten zu berechnen, da der Erfolg der Detektion der sichtbaren Beeren und Trauben indirekt über die den Erfolg der Ertragsschätzung evaluiert wird.

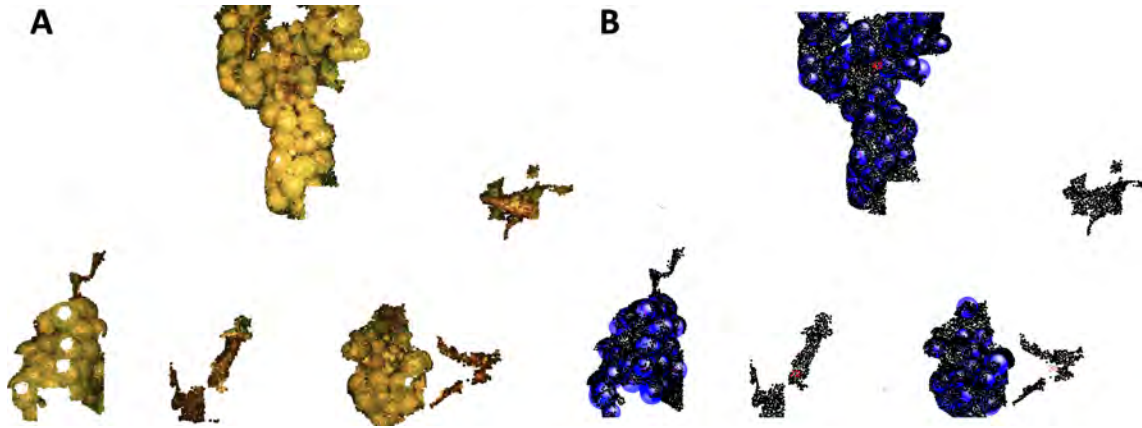


Abbildung 5.22: A: Traubenkandidaten aus Abschnitt 3 der MSS-Anbaureihe. Einige fehlklassifizierte Reste von Ästen zählen ebenfalls zu den Traubenkandidaten. B: Mit dem Berryfinder detektierte Beeren im dargestellten Ausschnitt. Traubenkandidaten, die keine Trauben, sondern Äste darstellen, wurden korrekt vom Berryfinder abgelehnt, weil bei ihnen weniger als $\min_B$ Beeren detektiert wurden.

	TS				MSS			
	$\mathcal{N}^T$	$\hat{\mathcal{E}}^T$	Δ	$\Delta[\%]$	$\mathcal{N}^T$	$\hat{\mathcal{E}}^T$	Δ	$\Delta[\%]$
Meter 1	32	30	-2	-6	61	72	10	16
Meter 2	25	31	6	24	47	39	-9	-20
Meter 3	33	35	2	6	53	77	24	45
Meter 4	22	24	2	9	48	50	2	4
Meter 5	21	27	6	28	68	44	-24	-35

Tabelle 5.20: Gesamtanzahl $\mathcal{N}^T$, sowie die geschätzte Gesamtanzahl $\hat{\mathcal{E}}^T$ und ihre absolute und prozentuale Abweichungen pro Abschnitt. Die horizontale Linie zwischen den Abschnitten 1 – 3 und 4 – 5 trennt unabhängige Ergebnisse von korrelierten Ergebnissen.

Tabelle 5.20 zeigt die tatsächliche Traubenanzahl im Vergleich zur geschätzten Traubenanzahl. Im Fall des ersten Abschnitts der TS-Anbaureihe werden zwei Trauben (6%) zu wenig geschätzt. In den übrigen vier Abschnitten wird die Traubenanzahl um bis zu 6 Trauben oder 28% höher als die tatsächliche Anzahl der Trauben geschätzt. Die Traubenanzahl der MSS-Anbaureihe wird teilweise bis zu 24 Trauben unterschätzt, teilweise aber auch bis zu 24 überschätzt. In Prozent beträgt die Unterschätzung bis zu 35%, die Überschätzung bis zu 45%.

	TS			MSS		
	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$	$\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)[g]$	$\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B)[g]$	$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$	$\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)[g]$	$\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B)[g]$
Meter 1	3248,0	683,8	0	426,0	335,8	2640,3
Meter 2	3988,6	709,1	0	264,0	180,7	1421,7
Meter 3	3887,1	785,1	0	530,0	361,6	2843,3
Meter 4	3084,4	531,8	0	271,0	232,4	1827,8
Meter 5	2842,2	607,8	0	336,0	206,6	1624,8

Tabelle 5.21: Gewichtsanteile $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)[g]$, $\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)[g]$ und $\hat{\mathcal{Y}}(\hat{\mathcal{O}}^B)[g]$ des Gesamtertrages $\hat{\mathcal{Y}}[g]$ jedes Abschnitts. Die Linie zwischen den Abschnitten 1 – 3 und 4 – 5 trennt unabhängige Ergebnisse von korrelierten Ergebnissen.

Tabelle 5.21 listet die einzelnen Gewichtsanteile am jeweiligen Gesamtertrag $\hat{\mathcal{Y}}[g]$ des Abschnitts nach den Formeln 5.29-5.31 auf. Die in Tabelle 5.21 gelisteten Teilerträge werden nach Formel 5.32 summiert und der geschätzte Ertrag $\hat{\mathcal{Y}}$ erhalten. Tabelle 5.22 listet den gewogenen Ertrag $\mathcal{Y}^A[g]$ der einzelnen Abschnitte und den geschätzten Ertrag $\hat{\mathcal{Y}}[g]$ jedes Abschnitts vergleichend auf.

	TS				MSS			
	$\mathcal{Y}^A[g]$	$\hat{\mathcal{Y}}[g]$	$\Delta[g]$	$\Delta[\%]$	$\mathcal{Y}^A[g]$	$\hat{\mathcal{Y}}[g]$	$\Delta[g]$	$\Delta[\%]$
Meter 1	4604,1	3931,8	-672,3	-15	2546,5	3402,1	855,6	34
Meter 2	4612,6	4697,7	85,1	2	2267	1866,4	-400,6	-18
Meter 3	4280,8	4672,9	392,1	9	3458,6	3734,9	-276,3	8
Meter 4	3163,7	3616,0	452,3	14	1707	2331,2	624,2	37
Meter 5	3589,2	3449,8	-139,4	-4	4004,3	2167,4	-1836,9	-46
σ			410,3	10			1082,4	35
$\overline{Meter}$			23,6	1			96,3	3
$\sum_{i=1}^5 Meter_i$	20250,4	20368,2	117,8	0,6	13983,4	13502,0	-481,4	-3

Tabelle 5.22: Tatsächlicher Ertrag $\mathcal{Y}^A[g]$, geschätzter Ertrag $\hat{\mathcal{Y}}[g]$ und ihre Abweichungen in $[g]$ und $[\%]$, sowie die jeweilige Standardabweichung, der Mittelwert, die Summe der Erträge aller Abschnitte sowie die Abweichung der Summen. Die Linie zwischen den Abschnitten 1 – 3 und 4 – 5 trennt unabhängige Ergebnisse von korrelierten Ergebnissen.

Die Einzelerträge der TS-Anbaureihe werden in zwei Fällen niedriger, in drei Fällen größer als der tatsächliche Ertrag geschätzt. Die geschätzten Erträge $\hat{\mathcal{Y}}[g]$ der einzelnen Abschnitte besitzen eine Standardabweichung von 10%.

In Gewichtseinheiten beträgt die Standardabweichung 410,3g. Der Mittelwert beträgt 23,6g bzw. 1%. Die Summe der geschätzten Einzelerträge der fünf Abschnitte weicht um 117,8g bzw. -0,6% von der Summe der tatsächlichen Einzelerträge ab.

Die Einzelerträge der MSS-Anbaureihe werden in zwei Fällen niedriger, in drei Fällen größer als der tatsächliche Ertrag geschätzt. Die Abweichungen in Gewicht nehmen teilweise Größenordnungen von über 1800g an, die Abweichungen in Prozent liegen bis -46% unter oder bis zu 37% über dem tatsächlichen Ertrag. Die Standardabweichung beträgt 35% bzw. 1082,4g, der Mittelwert 96,3g bzw. 3%.

Die Summe der geschätzten Einzelerträge der fünf Abschnitte weicht um -481,4g bzw. -3% von der Summe der tatsächlichen Einzelerträge ab.

Die Erträge der ersten drei Abschnitte sind nicht unabhängig von den bestimmten Kalibrierwerten. Die Ergebnisse sind daher zu optimistisch, jedoch sollte ein Eindruck von der Streuung der Abweichungen erhalten werden. Um eine von den Kalibrierwerten unabhängige Schätzung des Ertrages zu erhalten, werden nun ausschließlich die Erträge der Abschnitte 4 und 5 summiert und ihre Abweichung berechnet. Tabelle 5.23 listet die Ergebnisse auf. Für die TS-Anbaureihe weicht die Summe beider Erträge um 5% oder 312,9g voneinander ab. Die Summe beider Erträge weicht für die MSS-Anbaureihe um -21% oder -1212,7g voneinander ab.

	TS				MSS			
	$\mathcal{Y}[g]$	$\hat{\mathcal{Y}}[g]$	$\Delta[g]$	$\Delta[\%]$	$\mathcal{Y}[g]$	$\hat{\mathcal{Y}}[g]$	$\Delta[g]$	$\Delta[\%]$
$\sum_{i=4}^5 Meter_i$	6752,9	7065,8	312,9	5	5711,3	4498,6	-1212,7	-21

Tabelle 5.23: Summe der Erträge der Abschnitte 4 und 5 aus Tabelle 5.22, sowie ihre Abweichungen in $[g]$ und $[\%]$.

5.3.6 Diskussion: Ertragsschätzung im TS- und MSS-Erziehungssystem im BBCH89 Reifestadium

Zeitaufwand für die Ertragsschätzung

In einem Abschnitt von 1 Meter befinden sich 1 – 1,5 Millionen Punkte. Die Berechnung der SFHc-Deskriptoren benötigt für diese Punktzahl etwa 10 – 15 Minuten. Diese belegen 0,7 – 1,4Gb an Speicher. Die Klassifikation benötigt für diese Punktzahl rund 5 – 10 Minuten. Das CC-Verfahren arbeitet sehr schnell und benötigt weniger als 20 Sekunden für die Erstellung der einzelnen Komponenten. Der Berryfinder

benötigt für die Detektion der Beeren in einem Abschnitt von 1 Meter bis zu 15 Minuten. Für einen 1 Meter Abschnitt werden somit ungefähr 45 Minuten für die Schätzung der Ertragsparameter benötigt. Der gesamte Zeitaufwand für die Schätzung der Ertragsparameter einer 25m langen Anbaureihe beläuft sich demnach, inklusive der Rekonstruktionszeit von ungefähr 22[h], damit auf:

$$22[h] + (25 \cdot 3/4[h]) = 41[h]. \quad (5.39)$$

Um ebenfalls im Hochdurchsatz den Ertrag abschätzen zu können, muss dieser enorme Zeitaufwand reduziert werden. Im Kapitel 6-Fazit und Ausblick werden dazu einige Lösungsansätze beschrieben.

Größe der Stichprobe

Die Anzahl der Referenzwerte sowohl für die Berechnung der Kalibrierwerte, als auch zur Evaluation der Ergebnisse ist mit 25 Referenztrauben und fünf Abschnitten mit insgesamt 133 Trauben für die TS- bzw. 277 Trauben für die MSS-Anbaureihe gering. Die primäre Aufgabe dieser Arbeit ist die Entwicklung einer Methode zur Schätzung der Ertragsparameter und des Ertrages. Für einen *proof-of-concept* und eine erste Genauigkeitsanalyse, welche die potentiell erreichbare Genauigkeit der Ertragsschätzung ergründet, ist die Menge der Referenzwerte ausreichend. Um der Variabilität von Weinreben gerecht zu werden, müssen künftig Untersuchungen mit einer vielfach größeren Anzahl von Referenzwerten für die Berechnung der Kalibrierwerte und für die Evaluation durchgeführt werden (s. weitere Diskussion). Aufgrund der dazu notwendigen Datenmenge sind dies jedoch Aufgaben, die über diese Arbeit hinausgehen.

Schätzung des Ertrages der TS-Anbaureihe

Aus der tatsächlichen Beerenanzahl von 12 Referenztrauben wurde ein Kalibrierwert f_B für die Bestimmung der Anzahl $\hat{\mathcal{E}}_V^B$ bestimmt und anhand der tatsächlichen Beerenanzahl von 13 Referenztrauben überprüft. Dabei wurde eine Abweichung von 7% von der tatsächlichen Beerenanzahl der 13 Referenztrauben festgestellt. Ein Vergleich des tatsächlichen Ertrages aller 13 Referenztrauben mit dem geschätzten Ertrag $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ aller 13 Referenztrauben ergab eine Abweichung von -5% (Tabelle 5.8). Es kann daher konstatiert werden, dass die Schätzung des Ertrages der für den Sensor sichtbaren und detektierten Trauben über die Beerenanzahl mit einer Abweichung im einstelligen Prozentbereich vom tatsächlichen Ertrag für die TS-Anbaureihen möglich ist.

Dieses Ergebnis ist an den Erfolg der Beerendetektion gebunden, der für das TS-Erziehungssystem hoch ist. Das Blattwerk der TS-Anbaureihe ist in Traubennähe ausgedünnt worden. Die Trauben liegen daher sehr frei und können häufig beobachtet werden. In der Folge sind die Trauben mit einem weitestgehend gleichbleibenden Vollständigkeitsgrad und die Beeren mit einem hohen Detailgrad rekonstruiert worden. Der Detektionserfolg der Beeren und Trauben ist daher größtenteils konstant. Dies ermöglicht es, einen für die meisten Trauben repräsentativen Kalibrierwert f_B zu berechnen und die Ertragsschätzung mit geringer Abweichung vom tatsächlichen Ertrag durchzuführen.

Anhand der Tabellen 5.21 und 5.22 lässt sich erkennen, dass der Gewichtsanteil $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ der Beeren anteilig bereits bis ungefähr 3/4 oder mehr an den tatsächlichen Ertrag $\mathcal{Y}^A$ heranreicht. Dies zeigt, dass für das TS-Erziehungssystem beinahe sämtliche sichtbare Trauben für die Ertragsschätzung genutzt werden können. Der Erfolg der Ertragsschätzung für den gesamten Abschnitt hängt somit primär vom Erfolg der Beerendetektion in den sichtbaren Trauben und der Repräsentativität von f_B und g_B ab.

$\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ verbleibt stets kleiner als $\mathcal{Y}^A$, da einige Trauben nicht detektiert wurden. Durch $\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)$ soll dieser Unterschied abgeschätzt werden. Seine korrekte Abschätzung hängt im Gegensatz zu $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ von den geschätzten Traubenanzahlen $\hat{\mathcal{D}}_c^T$ und $\hat{\mathcal{M}}_l^T$ ab. Da sich die Ertragsschätzung somit sowohl aus einer Abschätzung der sichtbaren Beeren, als auch der Trauben zusammensetzt, korrespondiert eine Überschätzung der Traubenanzahl nicht zwangsläufig mit einer Überschätzung des Ertrages. Dieser Umstand zeigt sich in der Betrachtung der einzelnen Abschnitte:

- In den TS-Anbaureihen ist die geschätzte Anzahl der Trauben $\hat{\mathcal{E}}^T$ in den Abschnitten 2 – 5 um 2-6 Trauben oder bis zu 28% höher als die tatsächliche Anzahl der Trauben $\mathcal{N}^T$ (Tabelle 5.20). Die Überschätzung resultiert zum einen Teil aus der geschätzten Anzahl $\hat{\mathcal{M}}_l^T$ (Tabelle 5.19). Zum anderen Teil

resultiert sie aus Komponenten in $\mathcal{D}^T$, die nicht tatsächlichen Trauben entsprechen (False Positives). In diesen Komponenten werden ebenfalls Kugeln detektiert, die zu einer Überschätzung von Beeren und folglich des Ertrages beitragen können. Teilweise ist bereits $\hat{\mathcal{D}}_c^T$ höher als $\mathcal{N}^T$ (Abschnitt 2 und 5). Dennoch wird nur in den Abschnitten 2 – 4 der Ertrag überschätzt.

- Im Abschnitt 1 wurde sowohl die Anzahl der Trauben wie auch der geschätzte Ertrag im Vergleich zum tatsächlichen Ertrag zu niedrig geschätzt. Insgesamt wurden in diesem Abschnitt sieben tatsächliche Trauben von der Auswertemethode entfernt und konnten daher nicht zur Ertragsschätzung genutzt werden (Tabelle 5.14). Die Schätzung $\hat{\mathcal{M}}_l^T$ konnte diese Anzahl nur bedingt kompensiert werden: nur 3 der 7 tatsächlich fehlenden Trauben werden über f_m für die Ertragsschätzung hinzugenommen (Tabelle 5.19).
- Im Abschnitt 5 hingegen wird der Ertrag unterschätzt, obwohl die Anzahl der Trauben zu groß geschätzt wird. Dies ist ein weiteres Indiz dafür, dass der Erfolg der Ertragsschätzung im TS-Erziehungssystem primär vom Erfolg der Beerendetektion abhängt. Steht nur ein kleiner Teil einer Traube für die Beerendetektion zu Verfügung, bspw. weil sie nicht vollständig rekonstruiert oder ein Teil falsch klassifiziert wurde, können in dieser Traube zwangsläufig weniger Beeren detektiert werden. Tritt dies für den betrachteten Traubenbestand vermehrt auf, kann dies zu einer Unterschätzung des Ertrages führen. Die nur zum Teil rekonstruierte Traube wird dennoch als einzelne Traube gezählt und trägt zu $\hat{\mathcal{E}}^T$ bei, weshalb eine Überschätzung der Traubenanzahl bei gleichzeitiger Unterschätzung des Ertrages möglich ist.

Die geschätzten Erträge der fünf Abschnitte weichen zwischen -15% und 14% vom tatsächlichen Ertrag ab, bei einer Standardabweichung von 10% und einem Mittelwert von $1,2\%$. Da die Kalibrierwerte aus den ersten drei Abschnitten berechnet wurden, ist die tatsächliche Genauigkeit jedoch etwas schlechter.

In der Summe weichen sowohl die geschätzten Erträge der 13 Referenztrauben als auch die der fünf Abschnitte im einstelligen Prozentbereich vom tatsächlichen Ertrag ab (Tabellen 5.7, 5.22 und 5.23). Es sei nochmal betont, dass die Stichprobe klein ist. Ob diese Ergebnisse repräsentativ sind und wie groß die Stichprobe für die Berechnung der Kalibrierwerte sein sollte, müssen künftige Untersuchungen zeigen. Insbesondere sollten die Kalibrierwerte anhand einer größeren Anzahl unabhängiger Referenzwerte für den Ertrag überprüft und gegebenenfalls für ein optimales Ergebnis angepasst werden. Ein Anhaltspunkt für die Größe der Stichproben findet sich bei Nuske et al. [2014a], die die individuellen Abweichungen bei einer Stichprobe mehrerer Hundert Trauben im TS-Erziehungssystem untersucht haben und eine gleichmäßige Verteilung der individuellen Abweichungen konstatieren, die mit zunehmender Beobachtungszahl abnehmen. Ihre Ergebnisse könnten einen Startpunkt für weitere Untersuchungen der potentiell erreichbaren Genauigkeit der Ertragsschätzung im TS-Erziehungssystem geben.

Schätzung des Ertrages der MSS-Anbaureihe

Für die MSS-Anbaureihe ergab sich für die genutzten 10 Referenztrauben eine Abweichung von der tatsächlichen Beerenanzahl von -22% . Der Vergleich von $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ und tatsächlichem Ertrag aller 10 Referenztrauben ergab eine Abweichung von 33% (Tabelle 5.11). Die im Vergleich zur TS-Anbaureihe größere Abweichung erklärt sich 1. mit der durch Überbelichtung und Verdeckungsgrad verursachten schlechteren Rekonstruktion der Trauben. Recall und Precision von $\mathcal{D}^B$ werden daher eine größere Standardabweichung als bei der TS-Anbaureihe besitzen. Ein 2. Grund ist die unterschiedliche Vollständigkeit der rekonstruierten Trauben. Ist der Grad der Vollständigkeit aller rekonstruierten Trauben ungefähr konstant, dann können deren nicht detektierte und unsichtbare Beeren über f_B kompensiert werden. Durch die Verdeckung der Trauben ist dies in den MSS-Abschnitten nicht der Fall, der Vollständigkeitsgrad variiert. Der bestimmte Wert für f_B ist daher nur bedingt repräsentativ für alle Trauben.

Tabelle 5.18 zeigt, dass die Anzahl $\mathcal{O}^T$ aufgrund der hohen Dichte des Blattwerks im Vergleich zu $\hat{\mathcal{V}}^T$ sehr viel größer ist. Dementsprechend ist auch die geschätzte Anzahl $\hat{\mathcal{O}}^T$ deutlich größer, als $\hat{\mathcal{V}}^T$ (Tabelle 5.20). Dies hat Auswirkungen auf die Ertragsschätzung. Anhand der Tabellen 5.21 und 5.22 lässt sich erkennen, dass nur ein geringer Anteil des geschätzten Ertrages von den Gewichtsanteilen des sichtbaren Teils der Abschnitte $\hat{\mathcal{Y}}(\hat{\mathcal{E}}_V^B)$ und $\hat{\mathcal{Y}}(\hat{\mathcal{M}}_l^B)$ ausgemacht wird. Im Gegensatz zum TS-Erziehungssystem basiert die Ertragsschätzung im MSS-Erziehungssystem damit *nicht* auf einer direkten Detektion sichtbarer Beeren. Die Genauigkeit der

Ertragsschätzung hängt stattdessen primär davon ab, ob $\hat{O}^T$ und deren Gewicht $\hat{Y}(\hat{O}^B)$ korrekt abgeschätzt wird. Weinreben wachsen sehr variabel und eine Abschätzung der Anzahl stets verdeckter Trauben, ihrer Beeren über B_T und ihres Gewichts ist entsprechend unsicher, da die Kalibrierwerte die Varianz nur bedingt abbilden können. Dies ist in den Ergebnissen der fünf Abschnitte zu erkennen.

Die Tabellen 5.20 und 5.22 zeigen, dass es teilweise zu einer Unter- und teilweise zu einer Überschätzung der Traubenanzahl pro Abschnitt kommt und eine Unter- bzw. Überschätzung der Traubenanzahl mit einer Unter- bzw. Überschätzung des Ertrages korrespondiert. Die individuellen Abweichungen vom Ertrag sind im Vergleich zur TS-Anbaureihe deutlich größer, die Standardabweichung beträgt 35% (TS-Anbaureihe 10%). Dies ist auf die hohe Anzahl unsichtbarer Trauben und die Variabilität der Weinreben zurückzuführen. Letztere lässt sich bspw. anhand der Betrachtung der Traubenanzahl und des geschätzten Ertrages beobachten. Eine Überschätzung von lediglich 2 Trauben resultiert in Abschnitt 4 in einer Überschätzung des Ertrages von 37%, wohingegen eine Unterschätzung um 24 Trauben in Abschnitt 5 in einer betragsmäßig ähnlichen Unterschätzung des Ertrages von -46% resultiert.

In der Summe weichen die geschätzten Erträge der fünf Abschnitte um -3% vom tatsächlichen Ertrag ab (Tabelle 5.22). Aufgrund der Korrelation der ersten drei Abschnitte mit den Kalibrierwerten sind diese Ergebnisse jedoch zu optimistisch.

Für die unabhängigen Abschnitte 4 – 5 ergab sich eine jeweilige Abweichungen von 37% und -46% der Ertragsschätzung vom tatsächlichen Ertrag. In der Summe ergab sich eine Abweichung von -21% vom tatsächlichen Ertrag (Tabelle 5.23). Aufgrund der bisherigen Ergebnisse muss die Ertragsschätzung beim MSS-Erziehungssystem im Vergleich zum TS-Erziehungssystem als unsicherer angesehen werden. Um die größere Variabilität der Weinreben im MSS-Erziehungssystem zu berücksichtigen, müssen künftige Untersuchungen zur Berechnung, Überprüfung und eventuellen Anpassung der Kalibrierwerte mit einer noch größeren unabhängigen Stichprobe durchgeführt werden, als beim TS-Erziehungssystem. Orientiert an den bisherigen Ergebnissen von Nuske et al. [2014a] oder Aquino et al. [2018] (s. weitere Diskussion), die ausschließlich TS-Erziehungssysteme untersuchten, müsste die nötige Stichprobe für das MSS-Erziehungssystem mindestens mehrere Hundert bis evtl. Tausende Trauben umfassen.

Vergleich mit anderen Methoden der Ertragsschätzung

Im Vergleich mit anderen Methoden, welche Sensorplattformen für die Ertragsschätzung im gleichen Reifestadium einsetzen, werden ähnliche Genauigkeiten der Ertragsschätzung bei einer vergleichsweise größeren Stichproben erreicht. Nuske et al. [2014b] untersuchten mehrere Hundert Trauben und schildern Abweichungen zwischen 6,5 – 11,7% für Weinreben der TS-Erziehung mit ebenfalls ausgedünntem Blattwerk. Aquino et al. [2018] schätzen den Ertrag für das BBCH75 Reifestadium der TS-Erziehung und erreichen Abweichungen von 3,5 – 13,1%, für insgesamt 90 Weinreben mit ausgedünntem Blattwerk. Font et al. [2015] erreichen Abweichungen vom tatsächlichen Ertrag von 16 – 17% bei 25 Weinreben, nutzen jedoch eine von Hand bediente Kamera zur diskriminativen Bildaufnahme von individuellen Weinreben und sind daher mit dieser Arbeit nicht direkt vergleichbar.

Literatur zur Schätzung des Ertrages für Weinreben mit hoher Blattdichte oder für Weinreben im MSS-Erziehungssystem mit einer hohen Anzahl unsichtbarer Trauben wurden nach bisheriger Erkenntnis nicht veröffentlicht. Die hier dargestellten Ergebnisse für die MSS-Anbaureihe stellen die ersten Ergebnisse für die Schätzung der Ertragsparameter und des Ertrages unter den erschwerten Bedingungen einer hohen Blattdichte mit hohem Verdeckungsgrad dar und liefern wichtige Erkenntnisse für die künftige Entwicklung von Methoden der visuellen Ertragsschätzung. Verdeckungen lassen im Weinbau niemals gänzlich verhindern oder nur mit entsprechend hohem Aufwand durch Ausdünnung des Blattwerks reduzieren. Mit zunehmenden Verdeckungsgrad stoßen visuelle Methoden zur Ertragsschätzung zunehmend an ihre Grenzen und die Ertragsschätzung wird zunehmend unsicherer. Die Anforderungen an die Genauigkeit der Ertragsschätzung muss daher gegenüber dem Aufwand von Maßnahmen zur Erhöhung der Sichtbarkeit von Trauben abgewogen werden.

6. Fazit und Ausblick

6.1 Fazit

In dieser Arbeit wurde im Zuge des Projektes NoViSys folgende Aufgabe gestellt:

Entwicklung einer nicht-invasiven, sensorgestützten Hochdurchsatz-Methode zur automatischen Schätzung der Ertragsparameter und eine auf den Ertragsparametern basierende Schätzung des Ertrages für TS- und MSS-Erziehungssysteme im Reifestadium BBCH75 und BBCH89 unter Feldbedingungen.

Zur Lösung dieser Aufgaben wurden drei **Teilaufgaben zur Quantifikation der Ertragsparameter und des Ertrages** formuliert:

1. **Datenerhebung: Entwicklung und Untersuchung mobiler Sensorplattformen für die Datenerhebung im Hochdurchsatz unter Feldbedingungen für beide Erziehungssysteme**
Sensorplattformen für den Einsatz unter Feldbedingungen müssen robust und in der Lage sein, unabhängig von äußeren Umweltbedingungen Datenerhebungen im Hochdurchsatz durchzuführen. In dieser Arbeit wurden zwei bewegte Sensorplattformen auf ihre Eignung zur Lösung dieser Aufgabe untersucht:
 1) Der PHENObot ist ein Raupenfahrzeug mit einer RGB-Kamera und einem leistungsstarkem LED-Beleuchtungsrahmen, der direkt um die Kamera montiert ist. Das Kamerasystem in Kombination mit dem Beleuchtungsrahmen erzeugt gleichmäßig ausgeleuchtete, scharfe und detailreiche Bilder der Anbaureihen ohne Totalreflexionen und Schlagschatten. Er ist mit GPS zur Georeferenzierung der Bilder ausgestattet, jedoch wird die gemessene GPS-Position nicht zentimetergenau auf die tatsächliche Kameraposition transformiert. Aufgrund der Ausstattung mit nur einer RGB-Kamera, einer nur beschränkten Fahrgeschwindigkeit und keiner Abschirmung der Kamera vor äußeren Umweltbedingungen ist der PHENObot, trotz einem hoch technischen Kamera- und Beleuchtungssystem, nicht für die Datenerhebung im Hochdurchsatz unter Feldbedingungen geeignet.

 2) Der Phenoliner ist ein modifizierter Vollernter, der mit sechs Halogenstrahlern, einem Multi-Kamera-System (MKS) mit drei vertikal angeordneten RGB-Kameras, sowie GPS, einem Inklinometer und schnellen Speichermedien ausgestattet ist. Das MKS ist in dem verdunkelten Tunnel des Vollernters montiert, wodurch unter Einsatz der Beleuchtungseinheiten eine von äußeren Umweltbedingungen weitestgehend unabhängige Datenaufnahme ermöglicht wird. Die Bilder sind scharf und detailliert, die Beleuchtung ist jedoch stellenweise zu intensiv. Durch die Halogenstrahler werden Bilder mit einem Gelbstich erzeugt. Das Erziehungssystem hat Einfluss auf die Bildqualität hinsichtlich möglicher Überbelichtung und Verdeckungen der Kamera. Die Intensität der Halogenstrahler kann zu Überbelichtungen führen, falls Objekte eine zu kurze Distanz zu den Halogenstrahlern besitzen, wovon primär das MSS-Erziehungssystem betroffen ist. Die dem MKS gegenüberliegende Tunnelwand wurde zur Unterscheidung von Vorder- und Hintergrund farblich homogenisiert. Die Verwendung eines weißen Hintergrundes erwies sich jedoch als nicht optimal für die Hintergrundentfernung. Alle Bilder des MKS sind mittels GPS und einem Inklinometer zur Berücksichtigung der Fahrzeugneigung bis auf 2cm Genauigkeit georeferenziert. Dadurch kann eine zentimetergenaue Organisation der Ergebnisse in einem GIS stattfinden. Aufgrund der Nutzung von drei Kameras und einer erhöhten Fahrgeschwindigkeit kann eine Anbaureihe innerhalb von 2 – 3 Minuten bildlich erfasst werden. Der Phenoliner ist somit für die Datenerhebung im Hochdurchsatz unter Feldbedingungen geeignet.

Der Phenoliner stellt eine deutliche Verbesserung gegenüber dem PHENObot bezüglich Automation, Georeferenzierung, Geschwindigkeit, Vollständigkeit und Unabhängigkeit der Datenerhebung dar. Der PHENObot weist, neben der geringen Größe, die ihm eine hohe Flexibilität innerhalb der Korridore

für die Bildaufnahme ermöglicht, jedoch insbesondere hinsichtlich der Beleuchtung Vorteile gegenüber dem Phenoliner auf. Die zeigt sich vor allem anhand der Eignung beider Sensorplattformen für die Aufnahme von MSS-Anbaureihen. Mit dem PHENObot trat eine starke Überbelichtung, ungleichmäßige Ausleuchtung oder durch Berührung induzierte Bewegungen der MSS-Weinreben bei der Datenerhebung nicht auf. Daraus ergeben sich für die Zukunft zwei Lektionen: Ersten sollte die beim PHENObot angewandte Beleuchtungsstrategie für den Phenoliner adaptiert werden. Zweitens muss für ein Weg gefunden werden, Anbaureihen des MSS-Erziehungssystems ohne Berührungen von Sensorplattform und Weinreben zu erfassen.

2. Datenklassifikation: Entwicklung einer geeigneten algorithmenbasierten *Klassifikationsstrategie*, um in den Daten Beeren und Trauben von Blättern, Ästen und anderen Bestandteilen automatisch zu unterscheiden

Die Schätzung der Ertragsparameter und des Ertrages bedingt die vorherige Klassifikation der ungeordneten 3D-Punkte in semantische Klassen, so dass die semantisch geordneten 3D-Punkte zur Schätzung der Ertragsparameter genutzt werden können. Zu diesem Zweck wurde eine zweistufige Klassifikationsstrategie entwickelt, die auf dem Einsatz von geometrischen und farblichen Deskriptoren mit dem überwachten Klassifikator Import Vector Machine und der Graph Cut Methode GCO beruht. In Voruntersuchungen zur Klassifikation wurden sechs Deskriptoren auf ihre Eignung für die Einteilung der 3D-Punkte in die Klassen **Traube** und **Rest** untersucht. Dabei erwiesen sich der SFHc-Deskriptor als derjenige, mit dem ein optimales Klassifikationsergebnis erzielbar ist. Werden mit GCO räumlicher Nachbarbeziehungen zwischen den 3D-Punkten ausgenutzt, verbessern sich Recall und Precision um mehrere Prozentpunkte im zweistelligen Bereich.

Die Klassifikationsstrategie detektiert mit einem Recall von $\geq 89\%$ den Großteil der 3D-Punkte, die Trauben abbilden. Mit dieser Rate konnten alle Trauben in der Punktwolke für die weitere Auswertung genutzt werden. Traubenpunkte werden von anderen Objekten mit einer Precision zwischen 32 – 73% unterschieden. Die Precision für das MSS-Erziehungssystem ist aufgrund der hohen Objektdichte stets geringer als für das TS-Erziehungssystem.

Die Untersuchungen zeigen, dass die Klassifikationsstrategie für eine Vielzahl unterschiedlicher Erscheinungsformen geeignet ist. So konnten sowohl Daten des PHENObots und des Phenoliners, sowie Weinreben in beiden Reifestadien und Erziehungssystemen erfolgreich klassifiziert werden. Die Klassifikationsstrategie ist damit für eine Vielzahl von Anwendungen geeignet, muss jedoch vorher auf diese durch die Auswahl geeigneter Trainingsdaten und Berechnungsparameter für die SFHc-Deskriptoren angepasst werden.

3. Quantifikation: Entwicklung einer algorithmenbasierten Strategie zur automatischen *Schätzung der Ertragsparameter und des Ertrages* aus den klassifizierten Daten.

Die Schätzung der Ertragsparameter erfolgt anhand der als **Traube** klassifizierten 3D-Punkten. Anhand dieser Punkt werden die sichtbaren Trauben und Beeren detektiert. Ein Connected-Component Verfahren identifiziert dazu zusammenhängende Punktregionen, die als Traubenkandidaten an den Berryfinder zur Detektion von Beeren weiter gegeben werden. Im Berryfinder wird in den Traubenkandidaten nach Beeren gesucht und Traubenkandidaten verworfen, in denen eine Mindestanzahl an Beeren nicht detektiert werden konnte.

Die Detektion sichtbarer Beeren und Trauben wurde für die Reifestadien BBCH75 und BBCH89, sowie für die TS- und MSS-Erziehungssysteme durchgeführt. Es konnte gezeigt werden, dass durchschnittlich 60 – 72% aller Trauben detektiert werden. Die Anzahl nicht detektierter Trauben betrug dabei, unabhängig von Reifestadium oder Erziehungssystem, ungefähr ein Drittel der Anzahl sichtbarer Trauben. Die hohe Precision der Traubendetektion von 89 – 100% lässt den Schluss zu, dass beinahe alle als Traube akzeptierten Traubenkandidaten tatsächlichen Trauben entsprechen.

Im Schnitt werden 72 – 78% aller Beeren detektiert, bei einer durchschnittlichen Precision von 88 – 98%. Nicht rekonstruierte oder nicht korrekt klassifizierte Beeren, sowie Abweichungen der rekonstruierten Beeren von der Kugelgeometrie führen dazu, dass Beeren nicht detektiert werden können. Die hohe Precision zeigt, dass beinahe sämtliche Kugelschätzungen tatsächlichen Beeren entsprechen.

Die durchschnittliche geschätzte Beerengröße weicht unabhängig von Reifestadium und Erziehungssystem um 1 – 2,5mm von der durchschnittlichen Beerengröße aus Referenzmessungen ab. Die Abweichung

ist nach einem T-Test signifikant. In weiterführenden Untersuchungen muss die Ursache für die Abweichung identifiziert werden, in dem individuelle Beeren und ihre Größen untersucht werden.

Für fünf Abschnitte einer TS- und eine MSS-Anbaureihe im BBCH89 Reifestadium wurde eine Schätzung des Ertrages durchgeführt.

Die Ertragsschätzung für die TS-Anbaureihe erreichte Abweichungen vom tatsächlichen Ertrag von 5%. Potentiell kann die Schätzung des Ertrages für TS-Anbaureihen eine einstellige Abweichung vom tatsächlichen Ertrag erreichen, was durch die Rekonstruktion des Großteils der Trauben und den hohen Erfolg der Beerendetektion der sichtbaren Trauben gelingt. Die Beerendetektion erfolgt auf konstantem Niveau, da die Trauben gleichbleibend vollständig und die Beeren mit hohem Detailgrad rekonstruiert wurden. Dadurch kann ein für die meisten Trauben repräsentativer Kalibrierwert f_B berechnet werden. Wenn alle oder der Großteil der Trauben sichtbar sind, beruht der Erfolg der Ertragsschätzung auf der Detektion der sichtbaren Trauben und Beeren. Es sollte dann möglich sein, Erträge durch Vorher-Nachher Schätzungen zu vergleichen, bspw. um den Effekt der Ausdünnungsmaßnahmen zu bestimmen oder Verluste während der Saison zu quantifizieren. Künftige Untersuchungen mit größerer Stichprobenmenge müssen dies untersuchen.

Die Ertragsschätzung für die MSS-Anbaureihe erreichte Abweichungen vom tatsächlichen Ertrag von -21%. Im MSS-Erziehungssystem beruht durch den hohen Anteil stets verdeckter Trauben ein Großteil der Ertragsschätzung nicht auf der direkten Zählung sichtbarer Trauben und Beeren, sondern darauf, dass die Anzahlen der sensorisch nicht erfassbaren Beeren und Trauben genau abgeschätzt werden kann. Dazu müssen Kalibrierwerte und Mittelwerte verwendet werden, wodurch zwangsläufig Verzerrungseffekte mit höherem Einfluss auf die Ertragsschätzung als im TS-Erziehungssystem eingeführt werden. Mit der untersuchten Methode ist die Ertragsschätzung für das MSS-Erziehungssystem damit zwar möglich, jedoch nach bisherigen Erkenntnissen mit einer vergleichsweise größeren Abweichung vom tatsächlichen Ertrag. Ob Vergleiche der Erträge bei Vorher-Nachher Schätzungen für das MSS-Erziehungssystem in Anbetracht des hohen Anteils stets verdeckter Trauben sinnvoll sind, müssen weitere Untersuchungen mit größerer Stichprobenmenge zeigen (s. Ausblick).

Der Phenoliner und die entwickelte Methode zur Schätzung der Ertragsparameter und des Ertrages lösen die gestellten Teilaufgaben und ermöglichen eine Schätzung der Ertragsparameter und des Ertrages unter Erfüllung der in Kapitel 1 formulierten Kriterien: Die Datenerhebung erfolgt im Hochdurchsatz auf nicht-invasive Weise, geschieht im Nahbereich und erzeugt hochaufgelöste Daten. Die Daten sind durch die Verwendung von GPS metrisch und georeferenziert. Es konnte gezeigt werden, dass die Datenerhebung mit dem Phenoliner automatisiert bei standardisierten Aufnahmebedingungen geschieht. Die Einnahme einer Vielzahl von Perspektiven auf die Anbaureihen erfasst den Traubenbestand mit größtmöglicher Vollständigkeit. Die Datenklassifikation- und -quantifikation erfolgt automatisch und unter Berücksichtigung aktueller Referenzwerte.

Die im Rahmen von NoViSys formulierte Aufgabe kann für TS-Anbaureihen im BBCH89 Reifestadium grundsätzlich gelöst werden. Für das MSS-Erziehungssystem wurde ebenfalls eine Ertragsschätzung durchgeführt, jedoch müssen für dieses Erziehungssystem Maßnahmen zur Genauigkeitssteigerung unternommen werden (s. Ausblick). Im BBCH75 Reifestadium wurden für beide Erziehungssysteme erfolgreich die sichtbaren Beeren und Trauben detektiert, eine Ertragsschätzung für dieses Reifestadium bleibt noch durchzuführen (s. Ausblick).

6.2 Ausblick

Verbesserung der Ertragsschätzung

Bislang wurde das durchschnittliche Gewicht einer einzelnen Beere als Pauschalgewicht für die Ertragsschätzung verwendet. Um die Variabilität des Beerengewichts bei der Ertragsschätzung zu berücksichtigen, sollte in Zukunft der Durchmesser der Beeren zur Bestimmung des individuellen Gewichts einer Beere verwendet werden. Hierzu sind Untersuchungen notwendig, welche einen funktionalen Zusammenhang zwischen Beerengröße und Beerengewicht herstellt.

Des Weiteren sind die Ursachen für die Abweichungen zwischen tatsächlicher und geschätzter Beerengröße zu identifizieren. Dazu müssen individuelle Beeren manuell vermessen und ihre Größe mit dem Berryfinder geschätzt werden. Abweichungen könnten bspw. durch die manuelle Messung mit einem Messschieber entstehen, bei dem unterschiedliche Aufsatzpunkte oder die Krafteinwirkung des Anwenders zu unterschiedlichen Ergebnissen für die gleiche Beere führen kann. Eine weitere Ursache liegt in dem Umstand, dass Beeren von Natur aus nicht vollständig rund sind, der Berryfinder jedoch auf Kugelschätzungen basiert. Systematische Abweichungen aufgrund der gewählten Parameter des Berryfinders sind ebenfalls nicht auszuschließen.

Die Anzahl sichtbarer, aber nicht detektierter Trauben sollte verringert werden. Dazu muss primär die Precision der Klassifikation erhöht werden. Enthalten die Punktwolken weniger Regionen falsch klassifizierter Punkte, könnten die Schwellwerte p_K und min_B zur Filterung dieser verringert werden. Dadurch würden auch weniger Trauben fälschlicherweise entfernt. Dies könnte bspw. durch *Deep Learning* Klassifikation unter Verwendung der SFHc-Deskriptoren erreicht werden [Sun et al., 2017].

Vollständigere Erfassung des Traubenbestandes

Nur eine Seite der Anbaureihe wurde bildlich erfasst. Die Erfassung beider Seiten der Anbaureihen ist im Zuge einer vollständigeren Erfassung des Traubenbestandes erstrebenswert, da dadurch erhöhte Unabhängigkeit von Kalibrierwerten erlangt und die Genauigkeit potentiell gesteigert werden kann. Dies ließe sich einfach durch die erneute Aufnahme der Anbaureihe von der anderen Seite erreichen. Jedoch führt dies zu einer unnötigen Verdichtung des Bodens, was aus weinbaulicher Sicht vermieden werden sollte. Zudem verdoppelt sich auf diese Weise der zeitliche Aufwand für die Datenerhebung.

Eine technische Lösung besteht daher in einem MKS, welches Daten für beide Seiten der Anbaureihen gleichzeitig erhebt. Denkbar wäre an dieser Stelle die Montage von MKS auf beiden Seiten des Tunnels des Phenoliner. Um eine gegenseitige Aufnahme gegenüberliegend montierter Kameras zu vermeiden, müssten die MKS versetzt entlang der Längsachse des Fahrzeugs angebracht werden. Pro MKS würden in diesem Falle zwei separate Punktwolken erzeugt werden, da zwischen den Bildern beider MKS keine durchgehende Bildsequenz besteht. Eine Registrierung der Punktwolken über die GPS- und Inclinometer-Daten sollte jedoch Doppelzählungen des Traubenbestandes vermeiden können.

Ein weiterer Punkt ist der Vollständigkeitsgrad der bildlichen Abdeckung der gesamten Höhe der Weinreben durch das MKS im MSS-Erziehungssystem. Zwar erfasst das MKS einen Großteil des Traubenbestandes entlang der Höhe der Weinreben, eine weitere Kamera entlang der Vertikalen könnte die Bildabdeckung entlang der Vertikalen aber zusätzlich erhöhen. Dadurch würde jedoch auch die Datenmenge wachsen.

Adaption der Bildaufnahme auf das MSS-Erziehungssystem

Das geringe Platzangebot im Tunnel des Phenoliner und die voluminöse Ausprägung des Blattwerkes im MSS-Erziehungssystem stellt ein Problem für die Datenerhebung dar. Bei der Bildaufnahme mit dem Phenoliner kommt es durch die kürzere Distanz des MKS und der Beleuchtungseinheiten zu den Weinreben zur Überbelichtung und stellenweise zur Bewegung der Äste durch Kontakt mit dem Phenoliner. Dies erschwert die genaue Rekonstruktion der Weinreihen. Zur Erhöhung der Distanz kann das MKS tiefer im Phenoliner platziert werden. Eine größere Distanz der Beleuchtungseinheiten ist aufgrund des geringen Platzangebots im Tunnel jedoch nicht möglich. Zur Vermeidung von Berührungen von Weinreben und Phenoliner muss unter Umständen eine maschinelle Stützung des Blattwerks vor der Bildaufnahme in Betracht gezogen werden. Dies hätte zusammen mit der Betrachtung der Anbaureihen von zwei Seiten zudem den Vorteil, die Sichtbarkeit der Trauben weiter zu erhöhen.

Denkbar wäre außerdem die Verlagerung der Sensorik in einen gezogenen Anhänger, der mehr Platz für die Sensorik böte und ebenfalls von äußeren Einflüssen abgeschirmt werden kann. Dieser könnte breiter als der Tunnel des Phenoliner gebaut werden, um die Bewegungsfreiheit für die Sensorik zu erhöhen und den Abstand zu den Weinreben individuell optimieren zu können. Der Einsatz eines solchen Anhängers würde zudem die Verwendung von beidseitig angebrachten MKS ermöglichen (s. vorheriger Punkt).

Quantifikation der Ausdünnungsmaßnahmen für das MSS-Erziehungssystem

Es ist zu erwarten, dass nach der Ausdünnung durch Schütteln an den Weinreben Trauben bestehen bleiben, die stets verdeckt sind. Zusätzlich ist es möglich, dass ehemals verdeckte Trauben nach den Ausdünnungsmaßnahmen sichtbar und damit detektierbar werden. Für die Berechnung der Kalibrierwerte müssen die

nötigen Referenzwerte daher nach jeder Ausdünnung erneut bestimmt werden, da sich diese Anzahl erwartungsgemäß von den vorab ermittelten Anzahlen unterscheiden wird. Die Anzahl entfernter Trauben hängt jedoch von der Intensität und Dauer des Schüttelns ab. Für verschiedene Schüttelintensitäten müssten daher ebenfalls repräsentative Kalibrierwerte bestimmt werden, was die Komplexität der Ertragsschätzung zusätzlich erhöht und weitere Unsicherheiten einführt.

Eine Lösung wäre, den Ertrag vor der Ausdünnung visuell nach der hier beschriebenen Methode zu schätzen und das Gewicht der entfernten Trauben anschließend zu wiegen. Die entfernten Trauben könnten entweder direkt aufgefangen oder nachträglich gesammelt und gewogen werden. Der in der Anbaureihe verbliebene Ertrag wäre dann die Differenz von Anfangsertrag und dem Gewicht der entfernten Trauben. Auf diese Weise umginge man eine erneute visuelle und durch Kalibrierwerte verzerrte Ertragsschätzung zur Bestimmung des verbliebenen Ertrags der Anbaureihe.

Ertragsschätzung im BBCH75 Reifestadium

Ertragsschätzungen müssen ebenfalls im BBCH75 Reifestadium durchgeführt und anhand von Referenzmessungen evaluiert werden. Anhand von Recall und Precision, die in den Voruntersuchungen für die Schätzung der sichtbaren Beeren und Trauben im BBCH75 Reifestadium erreicht wurden, wird eine ähnliche Genauigkeit der Ertragsschätzung wie für das BBCH89 Reifestadium, in Abhängigkeit vom Erziehungssystem, erwartet.

Untersuchungen an Rotwein-Rebsorten

Bislang wurden lediglich Weißwein-Rebsorten untersucht. Das in dieser Arbeit beschriebene Vorgehen sollte künftig ebenfalls auf Rotwein-Rebsorten angewandt werden, um die Klassifikationsgenauigkeit zu evaluieren. Diago et al. [2012] und Dey et al. [2012] haben bereits Rotwein-Rebsorten erfolgreich klassifiziert. Es ist daher zu erwarten, dass die Klassifikation der Weinreben aufgrund der größeren farblichen Differenz des Blattwerks und der Trauben im Vergleich zu den grünlichen Trauben der Weißwein-Rebsorten mindestens genauso erfolgreich oder besser funktionieren wird.

Verbesserung der Beleuchtung und des Hintergrundes

Die Weinreben werden durch die Verwendung der Halogenstrahler grundsätzlich ausreichend hell ausgeleuchtet, wobei die Beleuchtung im Falle der MSS-Anbaureihen zu Überbelichtungen sehr naher Objekte führt. Um eine gleichmäßige Ausleuchtung zu erreichen und direkte Reflexionen zu vermeiden, sollten generell statt der punkthaften Halogenstrahler flächenhafte Lichtquellen mit diffuser Strahlung verwendet werden. Strategisch sollte sich dabei am Beleuchtungsrahmen des PHENObots orientiert werden, bei dem durchgehende Reihen dicht angebrachter LEDs mit Diffusoren und mit großem Öffnungswinkel verwendet wurden. Diffuse, flächenhafte Strahlung verhindert 1) das Entstehen von Schlagschatten und 2) das Auftreten von direkter Reflexion und erleuchtet die Szenerie mit gleichmäßiger Intensität. Wichtig ist auch die Anordnung der Beleuchtungseinheiten. Beim PHENObot ist der Beleuchtungsrahmen direkt um die Kamera montiert. Dadurch ist jedes einzelne Bild gleichmäßig ausgeleuchtet. Um eine gleichmäßige Ausleuchtung auch beim Phenoliner zu erreichen, sollte für das MKS möglichst dieselbe Anordnung der Beleuchtungseinheiten vorgenommen werden.

Das weiße Tuch sollte durch eine stabile Platte ersetzt werden, die sich während der Fahrt nicht bewegen kann. Deren Oberfläche sollte 1. wenig reflektieren und 2. von für Weinreben untypischer Farbe sein. Die Entfernung des Hintergrundes würde dadurch einfacher und die Traubendetektion präziser.

Beschleunigung und Automation der Auswertemethode

Durch den hohen Arbeitsaufwand der Ertragsschätzung entsteht ein neuer *Phänotypischer Flaschenhals* [Fiorani and Schurr, 2013; Minervini et al., 2015]). Um zeitnahe Maßnahmen für die Züchtung treffen zu können oder Entwicklungen zeitnah zu verfolgen, müssen die im Hochdurchsatz erhobenen Daten ebenfalls im Hochdurchsatz analysiert werden können. Die Integration von *High Throughput Analysis* Methoden in den Praxisalltag ist daher unerlässlich [Lussier and Liu, 2007; Shakoor et al., 2017]. Vorblickend ist damit zu rechnen, dass die exponentielle Entwicklung der Rechnertechnologien in den kommenden Dekaden auch diesen Flaschenhals für die Phänotypisierung überwinden wird. Eine Implementation der Auswertemethode auf parallelisierten Rechnern mit großem Arbeitsspeicher in einem Rechnernetz sollte dabei angestrebt werden.

Die vorgestellte Methode berechnet mehrere Zwischenergebnisse, die von mehreren Software Werkzeugen genutzt und weiterverarbeitet werden. Alle Teilschritte laufen automatisch ab, jedoch muss von einem Anwender zwischen den einzelnen Werkzeugen gewechselt und Zwischenergebnisse müssen manuell übergeben werden. Grundsätzlich sind alle einzelnen Methoden der Auswertemethode jedoch in einer einzigen Software implementierbar. Sie könnten in einer GUI (Graphical User Interface) übersichtlich für den Nutzer organisiert werden und die Ertragsschätzung nach Programmstart ohne manuelle Weitergabe der Zwischenergebnisse durch einen Anwender stattfinden.

Erzeugung der Punktwolken in Echtzeit

Ein weiterer interessanter Ansatz ist die Verwendung von Time-of-Flight Kameras wie der Kinect v2 (Microsoft Corp., Redmond, Washington, USA). Die Kinect v2 ist in der Lage, farbige 3D-Punktwolken von Objekten im Nahbereich in Echtzeit zu erzeugen [Kowalski et al., 2015] und wurde bereits im Weinbau für die Rekonstruktion von Trauben angewandt [Marinello et al., 2016]. Mehrere Kinect v2 Sensoren könnten vertikal angeordnet werden und aus der Fahrt heraus Punktwolken erzeugen. Letztere könnten durch vorherige Bestimmung der relativen Orientierung der Kinect v2 Sensoren zueinander registriert werden [Kowalski et al., 2015], um so die multiplen Perspektiven auszunutzen und Mehrfachzählungen der Beeren und Trauben zu vermeiden. Dieses Vorgehen würde durch Vermeidung der langen Berechnungszeit der Punktwolken im Postprocessing die Ertragsschätzung beschleunigen. Die Auflösung der Punktwolke hängt von der Entfernung zum Objekt ab, welche so groß gewählt werden muss, dass es nicht zu Berührungen mit den Weinreben kommt. Zu klären ist, ob die Auflösung der Punktwolken bei einer Mindestentfernung für die Ertragsschätzung ausreicht.

Weitere Einsatzmöglichkeiten der Auswertemethode und des Phenoliners

Über diese Arbeit hinausgehende Überlegungen betreffen die Ableitung weiterer phänotypischer Parameter aus den 3D-Punktwolken. Der große Vorteil der 3D-Rekonstruktion gegenüber von 2D-Methoden liegt in der dritten Dimension, die zusätzlich zur Farbe der Objekte erhalten wird. Informationen über die Geometrie der Pflanzenorgane, zu denen neben den Trauben auch Blätter, Stämme und Äste zählen, sind für Winzer von großem Interesse. Der *Leaf-Area-Index (LAI)* bspw. beschreibt die Fläche von Blättern und ist ein Indikator für die Gesundheit der Pflanze. Die Ausprägung des Laubwerks hat Einfluss auf das in der Weinrebe vorherrschende Mikroklima und wie intensiv Trauben dem Sonnenlicht ausgesetzt sind. In der Literatur existieren primär 2D-Methoden zur Schätzung des LAI (s. bspw. Fuentes et al. [2014] und Orlando et al. [2016]). Die 3D-Rekonstruktion der Weinreben in metrischen Einheiten ermöglicht hingegen die Bestimmung der individuellen Fläche pro Blatt über die rekonstruierten Geometrien bspw. durch *Delaunay-Triangulierung* oder *NURBS* [Piegl, 1991; Santos and Ueda, 2013], sowie Ableitung von Position oder Neigung der Blätter. Ebenfalls von Bedeutung ist das Gewicht des Holzes [Kicherer, 2015], welches von Länge und Umfang der Stämme/Äste abhängt. Der Umfang rekonstruierter Äste kann bspw. über Zylinderschätzungen [Paulus et al., 2014b; Rose et al., 2015] und die Länge des Astgeflechts bspw. über Methoden der *Skelettierung* [Yan et al., 2009] bestimmt werden.

Begriffe und Abkürzungen

Begriffsdefinitionen

An dieser Stelle seien einige weinbauliche Begriffe knapp definiert:

- Weinrebe (*vitis vinifera*): Kletterpflanze, deren Fruchtbestand in Trauben wächst.
- Traube: Frucht der Weinrebe. Ein Verbund mehrerer Weinbeeren an einem Stielgerüst.
- Weinbeere: Eine einzelne Beere. Mehrere Beeren konstituieren eine Traube.
- Stielgerüst: Holzartiges Gerüst, an dem die einzelnen Beeren wachsen.
- Anbaureihe: Mehrere Weinreben in einer Reihe von ca. 25m, in einem Abstand von ca. einem Meter gepflanzt.
- Blattwerk: Umfasst Blätter und Äste der Weinrebe.

Abkürzungen

BBCH Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie

CC Connected Components (Verfahren)

CloudC Cloud Compare

CSHOT Color-Signature-of-Histograms-of-Orientations

GCO Graph-Cut-Optimization

GSD Ground Sampling Distance

HSV Hue-Saturation-Value

IVM Import Vector Machine

MSS Minimalschnitt im Spalier Erziehungssystem

NoViSys Novel viticulture systems for sustainable production and products

PIWI Pilzwiderstandsfähige (Rebsorte)

RF Random Forest

SF(c) Saliency Feature (Color)

SFH(c) Surface Feature Histogram (Color)

SfM/MVS Structure from Motion/Multi-View-Stereo

SVM Support Vector Machine

TS Traditionelles Spalier Erziehungssystem

Literaturverzeichnis

- Aizerman, M. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, 25:821–837, 1964.
- Alexa, M.; Behr, J.; Cohen-Or, D.; Fleishman, S.; Levin, D., and Silva, C. T. Computing and rendering point set surfaces. *IEEE Transactions on visualization and computer graphics*, 9(1):3–15, 2003.
- Alhamzi, K.; Elmogy, M., and Barakat, S. 3D object recognition based on local and global features using point cloud library. *International Journal of Advancements in Computing Technology*, 7(3):43, 2015.
- Aquino, A.; Diago, M. P.; Millán, B., and Tardáguila, J. A new methodology for estimating the grapevine-berry number per cluster using image analysis. *Biosystems Engineering*, 156:80–95, 2017.
- Aquino, A.; Millan, B.; Diago, M.-P., and Tardaguila, J. Automated early yield prediction in vineyards from on-the-go image acquisition. *Computers and Electronics in Agriculture*, 144:26–36, 2018.
- Araus, J. L. and Cairns, J. E. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, 19(1):52–61, 2014.
- Behley, J.; Steinhage, V., and Cremers, A. B. Performance of histogram descriptors for the classification of 3D laser range data in urban environments. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 4391–4398, 2012.
- Benning, W. *Statistik in Geodäsie, Geoinformation und Bauwesen*. Herbert Wichman Verlag, Heidelberg, Germany, 4. edition, 2011.
- Bentley, J. L. K-d trees for semidynamic point sets. In *Proceedings of the sixth annual symposium on Computational geometry*, pages 187–197, 1990.
- Bora, D. J.; Gupta, A. K., and Khan, F. A. Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation. *International Journal of Emerging Technology and Advanced Engineering*, 5(2):192–203, 2015.
- Bostrom, H. Estimating class probabilities in random forests. In *Machine Learning and Applications, 2007. ICMLA 2007. Sixth International Conference on*, pages 211–216, 2007.
- Boykov, Y. and Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE transactions on pattern analysis and machine intelligence*, 26(9):1124–1137, 2004.
- Boykov, Y.; Veksler, O., and Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11):1222–1239, 2001.
- Braun, A. C.; Weidner, U., and Hinz, S. Support vector machines, import vector machines and relevance vector machines for hyperspectral classification - a comparison. In *Evolution in Remote Sensing (WHISPERS), 3rd Workshop on Hyperspectral Image and Signal Processing*, pages 1–4, 2011.
- Breiman, L. Random forests. *Machine learning*, 45(1):5–32, 2001.
- Chabbi, A. and Loescher, H. W. *Terrestrial Ecosystem Research Infrastructures: Challenges and Opportunities*. CRC Press, 2017.
- Chawla, N. V. and Cieslak, D. A. Evaluating probability estimates from decision trees. In *American Association for Artificial Intelligence*, 2006.
- Clingeffer, P. R.; Martin, S.; Krstic, M., and Dunn, G. M. *Crop Development, Crop Estimation and Crop Control to Secure Quality and Production of Major Wine Grape Varieties: A National Approach: Final Report to Grape and Wine Research & Development Corporation*. Grape and Wine Research & Development Corporation, 2001.
- Commission Internationale De l'Eclairage. CIELab. <http://www.cie.co.at/>, 2017. [Online; accessed 2017-10-17].
- Condurso, C.; Cincotta, F.; Tripodi, G.; Sparacio, A.; Giglio, D. M. L.; Sparla, S., and Verzera, A. Effects of cluster thinning on wine quality of syrah cultivar (*vitis vinifera* l.). *European Food Research and Technology*, 242(10):1719–1726, 2016.

- Cortes, C. and Vapnik, V. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- Danielgm. CloudCompare. <http://www.danielgm.net/cc/>, 2017. [Online; accessed 2017-09-12].
- Delong, A. GCO. <http://www.psi.toronto.edu/~andrew/>, 2016. [Online; accessed 2017-08-12].
- Delong, A.; Osokin, A.; Isack, H. N., and Boykov, Y. Fast approximate energy minimization with label cost. *International Journal of Computer Vision*, 96(1):1–27, 2012. doi: 10.1007/s11263-011-0437-z.
- Dey, D.; Mummert, L., and Sukthankar, R. Classification of plant structures from uncalibrated image sequences. *Proceedings of IEEE Workshop on Applications of Computer Vision*, pages 329–336, 2012. ISSN 21583978. doi: 10.1109/WACV.2012.6163017.
- Dhondt, S.; Wuyts, N., and Inzé, D. Cell to whole-plant phenotyping: the best is yet to come. *Trends in Plant Science*, 18(8):433–444, 2013. ISSN 1360-1385.
- Diago, M.-P.; Correa, C.; Millán, B.; Barreiro, P.; Valero, C., and Tardaguila, J. Grapevine yield and leaf area estimation using supervised classification methodology on RGB images taken under field conditions. *MDPI Sensors*, 12(12):16988–7006, 2012. ISSN 1424-8220. doi: 10.3390/s121216988.
- Diago, M. P.; Tardaguila, J.; Aleixos, N.; Millan, B.; Prats-Montalbán, J. M.; Cubero, S., and Blasco, J. Assessment of cluster yield components by image analysis. *Journal of the Science of Food and Agriculture*, 95(6):1274–1282, 2015.
- Duda, R. O. and Hart, P. E. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.
- Duda, R. O.; Hart, P. E., and Stork, D. G. *Pattern classification*. John Wiley & Sons, 2012.
- Dupuis, J. Genauigkeitsaspekte bei Messsystemen mit Nahbereichslaserscannern zum Pflanzenmonitoring. Dissertation, Universitäts-und Landesbibliothek Bonn, 2017.
- Fernández, R.; Montes, H.; Salinas, C.; Sarria, J., and Armada, M. Combination of rgb and multispectral imagery for discrimination of cabernet sauvignon grapevine elements. *MDPI Sensors*, 13(6):7838–7859, 2013.
- Fiorani, F. and Schurr, U. Future scenarios for plant phenotyping. *Annual review of plant biology*, 64: 267–291, 2013.
- Font, D.; Tresanchez, M.; Martínez, D.; Moreno, J.; Clotet, E., and Palacín, J. Vineyard yield estimation based on the analysis of high resolution images obtained with artificial illumination at night. *MD-PI Sensors*, 15(4):8284–8301, 2015. ISSN 1424-8220. doi: 10.3390/s150408284.
- Franc, V.; Zien, A., and Schölkopf, B. Support vector machines as probabilistic models. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 665–672, 2011.
- Fuentes, S.; Poblete-Echeverría, C.; Ortega-Farías, S.; Tyerman, S., and De Bei, R. Automated estimation of leaf area index from grapevine canopies using cover photography, video and computational analysis methods. *Australian journal of grape and wine research*, 20(3):465–473, 2014.
- Furbank, R. T. and Tester, M. Phenomics—technologies to relieve the phenotyping bottleneck. *Trends in plant science*, 16(12):635–44, 2011. ISSN 1878-4372. doi: 10.1016/j.tplants.2011.09.005.
- Furukawa, Y.; Hernández, C., et al. Multi-view stereo: A tutorial. *Foundations and Trends® in Computer Graphics and Vision*, 9(1-2):1–148, 2015.
- Hartley, R. I. and Zisserman, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- Herrero-Huerta, M.; Gonzalez-Aguilera, D.; Rodriguez-Gonzalez, P., and Hernandez-López, D. Vineyard yield estimation by automatic 3D bunch modelling in field conditions. *Computers and Electronics in Agriculture*, 110:17–26, 2015. ISSN 01681699. doi: 10.1016/j.compag.2014.10.003.
- Herzog, K.; Roscher, R.; Wieland, M.; Kicherer, A.; Läbe, T.; Förstner, W.; Kuhlmann, H., and Töpfer, R. Initial steps for high-throughput phenotyping in vineyards. *VITIS-Journal of Grapevine Research*, 53(1):1, 2014.
- Ivorra, E.; Sánchez, A.; Camarasa, J.; Diago, M. P., and Tardaguila, J. Assessment of grape cluster yield components based on 3D descriptors using stereo vision. *Food Control*, 50:273–282, 2015.
- Jain, A. K.; Duin, R. P. W., and Mao, J. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1):4–37, 2000.

- Kicherer, A. High-throughput phenotyping of yield parameters for modern grapevine breeding. Dissertation, Julius Kühn-Institut Siebeldingen, 2015.
- Kicherer, A.; Roscher, R.; Herzog, K.; Förstner, W., and Töpfer, R. Image based evaluation for the detection of cluster parameters in grapevine. In *XI International Conference on Grapevine Breeding and Genetics 1082*, pages 335–340, 2014.
- Kicherer, A.; Herzog, K.; Pflanz, M.; Wieland, M.; Rüger, P.; Kecke, S.; Kuhlmann, H., and Töpfer, R. An automated field phenotyping pipeline for application in grapevine research. *MDPI Sensors*, 15(3):4823–36, 2015a. ISSN 1424-8220. doi: 10.3390/s150304823.
- Kicherer, A.; Roscher, R.; Herzog, K.; Šimon, S.; Förstner, W., and Töpfer, R. BAT (Berry Analysis Tool): A high-throughput image interpretation tool to acquire the number, diameter, and volume of grapevine berries. *VITIS-Journal of Grapevine Research*, 52(3):129, 2015b.
- Kicherer, A.; Herzog, K.; Bendel, N.; Klück, H.-C.; Backhaus, A.; Wieland, M.; Rose, J. C.; Klingbeil, L.; Läbe, T.; Hohl, C., et al. Phenoliner: A new field phenotyping platform for grapevine research. *MDPI Sensors*, 17(7):1625, 2017.
- Klodt, M.; Herzog, K.; Töpfer, R., and Cremers, D. Field phenotyping of grapevine growth using dense stereo reconstruction. *BMC bioinformatics*, 16(1): 143, 2015.
- Knapitsch, A.; Park, J.; Zhou, Q.-Y., and Koltun, V. Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics (TOG)*, 36(4):78, 2017.
- Kolmogorov, V. and Zabini, R. What energy functions can be minimized via graph cuts? *IEEE transactions on pattern analysis and machine intelligence*, 26(2):147–159, 2004.
- Kotsiantis, S. B. Supervised machine learning: A review of classification techniques. *Informatica*, 31: 249–268, 2007. ISSN 09226389. doi: 10.1115/1.1559160.
- Kowalski, M.; Naruniec, J., and Daniluk, M. Live scan3D: A fast and inexpensive 3D data acquisition system for multiple kinect v2 sensors. In *IEEE International Conference on 3D Vision (3DV)*, pages 318–325, 2015.
- Kumar, N. Using Support Vector Machines Effectively. <http://neerajkumar.org/writings/svm/>, 2017. [Online; accessed 2017-08-14].
- Küng, O.; Strecha, C.; Beyeler, a.; Zufferey, J.-C.; Floreano, D.; Fua, P., and Gervais, F. The accuracy of automatic photogrammetric techniques on ultra-light UAV imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22:125–130, 2012. ISSN 1682-1777. doi: 10.5194/isprsarchives-XXXVIII-1-C22-125-2011.
- Lalonde, J.-F.; Vandapel, N.; Huber, D. F., and Hebert, M. Natural terrain classification using three-dimensional ladar data for ground robot mobility. *Journal of field robotics*, 23(10):839–861, 2006.
- Liu, S. and Whitty, M. Automatic grape bunch detection in vineyards with an svm classifier. *Journal of Applied Logic*, 13(4):643–653, 2015.
- Liu, S.; Whitty, M., and Cossell, S. A lightweight method for grape berry counting based on automated 3D bunch reconstruction from a single image. In *Workshop on Robotics in Agriculture, Seattle, USA*, 2015.
- Lloyd, S. Least squares quantization in PCM. *IEEE transactions on information theory*, 28(2):129–137, 1982.
- Lorenz, D.; Eichhorn, K.; Bleiholder, H.; Klose, R.; Meier, U., and Weber, E. Growth stages of the grapevine: Phenological growth stages of the grapevine (*Vitis vinifera* L. ssp. *vinifera*) codes and descriptions according to the extended bbch scale. *Australian Journal of Grape and Wine Research*, 1(2):100–103, 1995.
- Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- Lumia, R.; Shapiro, L., and Zuniga, O. A new connected components algorithm for virtual memory computers. *Computer Vision, Graphics, and Image Processing*, 22(2):287–300, 1983.
- Luo, L.; Tang, Y.; Zou, X.; Wang, C.; Zhang, P., and Feng, W. Robust grape cluster detection in a vineyard by combining the adaboost framework and multiple color components. *MDPI Sensors*, 16(12): 2098, 2016.
- Lussier, Y. A. and Liu, Y. Computational approaches to phenotyping: high-throughput phenomics. *Proceedings of the American Thoracic Society*, 4(1): 18–25, 2007.

- 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*, 135:300–311, 2017a.
- Mack, J.; Trakowski, A.; Rist, F.; Herzog, K., and Töpfer, R. Experimental evaluation of the performance of local shape descriptors for the classification of 3D data in precision farming. *Journal of Computer and Communications*, 5(12):1, 2017b.
- Mahalanobis, P. C. On the generalized distance in statistics. National Institute of Science of India, 1936.
- Marinello, F.; Pezzuolo, A.; Cillis, D., and Sartori, L. Kinect 3D reconstruction for quantification of grape bunches volume and mass. *Conference on Engineering for Rural Development, Jelgava, Latvia*, 15:876–881, 2016.
- Matese, A. and Di Gennaro, S. Technology in precision viticulture: A state of the art review. *Int. J. Wine Res*, 7:69–81, 2015.
- Matthews, M. and Nuzzo, V. Berry size and yield paradigms on grapes and wines quality. *Acta Horticulturae*, 754:423, 2007.
- Minervini, M.; Scharf, H., and Tsafaris, S. A. Image analysis: The new bottleneck in plant phenotyping [applications corner]. *IEEE signal processing magazine*, 32(4):126–131, 2015.
- Mirbod, O.; Yoder, L., and Nuske, S. Automated measurement of berry size in images. *IFAC-PapersOnLine*, 49(16):79–84, 2016.
- Naor, A.; Gal, Y., and Bravdo, B. Shoot and cluster thinning influence vegetative growth, fruit yield, and wine quality of sauvignon blanc grapevines. *Journal of the American Society for Horticultural Science*, 127(4):628–634, 2002.
- Niculescu-Mizil, A. and Caruana, R. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd international conference on Machine learning*, pages 625–632, 2005.
- NoViSys. Novel viticulture systems for sustainable production and products - BMBF: FKZ 031A349. <http://www.zukunft-weinbau.de/>, 2017. [Online; accessed 2017-11-04].
- Nuske, S.; Achar, S.; Bates, T.; Narasimhan, S., and Singh, S. Yield estimation in vineyards by visual grape detection. *IEEE International Conference on Intelligent Robots and Systems*, pages 2352–2358, 2011. ISSN 2153-0858. doi: 10.1109/IROS.2011.6048830.
- Nuske, S.; Gupta, K.; Narasimhan, S., and Singh, S. Modeling and calibrating visual yield estimates in vineyards. In *Field and Service Robotics*, pages 343–356, 2014a.
- Nuske, S.; Wilshusen, K.; Achar, S.; Yoder, L.; Narasimhan, S., and Singh, S. Automated visual yield estimation in vineyards. *Journal of Field Robotics*, 31(5):837–860, 2014b.
- Orlando, F.; Movedi, E.; Coduto, D.; Parisi, S.; Brancadoro, L.; Pagani, V.; Guarneri, T., and Confalonieri, R. Estimating leaf area index (lai) in vineyards using the pocketlai smart-app. *MDPI Sensors*, 16(12):2004, 2016.
- Patki, P. S. and Kelkar, V. V. Classification using different normalization techniques in support vector machine. In *International Conference on Communication Technology*, pages 17–19, 2013.
- Paulus, S.; Dupuis, J.; Mahlein, A.-K., and Kuhlmann, H. Surface feature based classification of plant organs from 3D laserscanned point clouds for plant phenotyping. *BMC Bioinformatics*, 14(1):238, 2013.
- Paulus, S.; Behmann, J.; Mahlein, A.-K.; Plümer, L., and Kuhlmann, H. Low-cost 3D systems: suitable tools for plant phenotyping. *MDPI Sensors*, 14(2):3001–18, 2014a. ISSN 1424-8220. doi: 10.3390/s140203001.
- Paulus, S.; Dupuis, J.; Riedel, S., and Kuhlmann, H. Automated analysis of barley organs using 3D laser scanning: an approach for high throughput phenotyping. *MDPI Sensors*, 14(7):12670–86, 2014b. ISSN 1424-8220. doi: 10.3390/s140712670.
- Penasa, L. CloudCompare PCL Plugin. <http://www.danielgm.net/cc/>, 2017. [Online; accessed 2017-09-12].
- PHENOVines. Hochdurchsatzphänotypisierung von Ertragsparametern und Mehлтаubefall bei Weinreben - BMBF: FKZ 0315968A/B. <http://www.hs-geisenheim.de/forschungszentren/institut-fuer-technik/forschung/phenovines.html>, 2017. [Online; accessed 2017-11-04].
- Piegl, L. On NURBS: A survey. *IEEE Computer Graphics and Applications*, 11(1):55–71, 1991.

- Platt, J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10 (3):61–74, 1999.
- Project CROP.SENSE. - Sensortechnology for plant breeding and management - BMBF: FKZ 0315534) . <https://www.cropsense.uni-bonn.de/>, 2017. [Online; accessed 2017-11-04].
- Rabatel, G. and Guizard, C. Grape berry calibration by computer vision using elliptical model fitting. In *Precision Agriculture 2007, Proceedings of the 6th European Conference on Precision Agriculture*, pages 581–587, 2007.
- Remondino, F.; Spera, M. G.; Nocerino, E.; Menna, F., and Nex, F. State of the art in high density image matching. *The Photogrammetric Record*, 29(146):144–166, 2014. ISSN 0031868X. doi: 10.1111/phor.12063.
- Reynolds, A.; Pool, R., and Matpick, L. Influence of cluster exposure on fruit composition and wine quality of seyval blanc grapes. *Vitis*, 25:85–95, 1986.
- Roscher, R. Sequential learning using Incremental Import Vector Machines for semantic segmentation. Dissertation, Universitäts-und Landesbibliothek Bonn, 2012.
- Roscher, R.; Förstner, W., and Waske, B. I2VM: Incremental import vector machines. *Image and Vision Computing*, 30(4-5):263–278, 2012a. ISSN 02628856. doi: 10.1016/j.imavis.2012.04.004.
- Roscher, R.; Waske, B., and Förstner, W. Evaluation of import vector machines for classifying hyperspectral data. *Department of Photogrammetry, University of Bonn*, 2012b.
- Roscher, R.; Herzog, K.; Kunkel, A.; Kicherer, A.; Töpfer, R., and Förstner, W. Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. *Computers and Electronics in Agriculture*, 100:148–158, 2014.
- Rose, J. C.; Paulus, S., and Kuhlmann, H. Accuracy analysis of a multi-view stereo approach for phenotyping of tomato plants at the organ level. *MDPI Sensors*, 15(5):9651–9665, 2015.
- Rose, J. C.; Kicherer, A.; Wieland, M.; Klingbeil, L.; Töpfer, R., and Kuhlmann, H. Towards automated large-scale 3D phenotyping of vineyards under field conditions. *MDPI Sensors*, 16(12):2136, 2016.
- Rusu, R. B. Semantic 3D object maps for everyday manipulation in human living environments. Dissertation, 2009.
- Santos, T. and Ueda, J. Automatic 3D plant reconstruction from photographs, segmentation and classification of leaves and internodes using clustering. In *International Conference on Function-Structural Plant Models*, number 7, pages 95–97, 2013. ISBN 9789516514089.
- Santos, T. T.; Bassoi, L. H.; Oldoni, H., and Martins, R. L. Automatic grape bunch detection in vineyards based on affordable 3D phenotyping using a consumer webcam. In *Conference: XI Congresso Brasileiro de Agroinformatica (SBIAgro), At Campinas, SP - Brazil*, 2017.
- Schöler, F. and Steinhage, V. Automated 3D reconstruction of grape cluster architecture from sensor data for efficient phenotyping. *Computers and Electronics in Agriculture*, 114:163–177, 2015.
- Shakoor, N.; Lee, S., and Mockler, T. C. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Current opinion in plant biology*, 38:184–192, 2017.
- Strecha, C.; Bronstein, A. M.; Bronstein, M. M., and Fua, P. Ldhash: Improved matching with smaller descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(1):66–78, 2012. ISSN 0162-8828. doi: <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.103>.
- Sun, Y.; Liu, Y.; Wang, G., and Zhang, H. Deep learning for plant identification in natural environment. *Computational intelligence and neuroscience*, 2017:1–6, 2017.
- Tombari, F.; Salti, S., and Di Stefano, L. Unique signatures of histograms for local surface description. In *European conference on computer vision*, pages 356–369, 2010.
- Tombari, F.; Salti, S., and Di Stefano, L. A combined texture-shape descriptor for enhanced 3D feature matching. In *18th IEEE International Conference on Image Processing (ICIP)*, pages 809–812, 2011.
- Töpfer, R.; Hausmann, L.; Harst, M.; Maul, E.; Zyprian, E., and Eibach, R. New horizons for grapevine breeding. *Methods in temperate fruit breeding. fruit, vegetable and cereal science and biotechnology*, 5:79–100, 2011.

- Tsingas, V. *Automatisierung der Punktübertragung in der Aerotriangulation durch mehrfache digitale Bildzuordnung*. Dissertation, Verlag der Bayerischen Akademie der Wissenschaften München, 1992.
- Vapnik, V. *The nature of statistical learning theory*. Springer Science & Business Media, 2013.
- Walg, O. Minimal pruning of trellis trained grapevines: spring activities. *Das Deutsche Weinmagazin*, 8:32–33, 2012a.
- Walg, O. Minimal pruning of trellis trained grapevines: thinning by grape harvester. *Das Deutsche Weinmagazin*, 13:12–15, 2012b.
- Walg, O. Minimalschnitt im Spalier: Erkenntnisse aus 2011. *Der Deutsche Weinbau*, 22:26–29, 2012c.
- Weinmann, M.; Schwartz, C.; Ruiters, R., and Klein, R. A multi-camera, multi-projector super-resolution framework for structured light. *International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pages 397–404, 2011.
- Whalley, J. and Shanmuganathan, S. Applications of image processing in viticulture: A review. In *20th MSSANZ-International Congress on Modeling and Simulation (MODSIM)*, pages 531–538, 2013.
- Willow Garage. Point Cloud Library. <http://www.pointclouds.org/>, 2017. [Online; accessed 2017-10-17].
- Witte, B. and Schmidt, H. *Vermessungskunde und Grundlagen der Statistik für das Bauwesen*. Herbert Wichman Verlag, Heidelberg, Germany, 6. edition, 2006.
- Yan, D.-M.; Wintz, J.; Mourrain, B.; Wang, W.; Boudon, F., and Godin, C. Efficient and robust reconstruction of botanical branching structure from laser scanned points. In *11th IEEE International Conference on Computer-Aided Design and Computer Graphics.*, pages 572–575, 2009.
- Zhu, J. and Hastie, T. Kernel Logistic Regression and the Import Vector Machine. *Journal of Computational and Graphical Statistics*, 14:185–205, 2005. ISSN 1061-8600. doi: 10.1198/106186005X25619.

Danksagung

Mein erster Dank gilt Prof. Dr. Heiner Kuhlmann, der mir die Möglichkeit gab, am IGG unter seiner Professur im NoViSys-Projekt diese Arbeit zu verfassen. Für seine Betreuung, Unterstützung und Ratschläge seien ihm sehr herzlich gedankt. Meinen Korreferenten Prof. Dr. Ribana Roscher und Prof. Dr. Reinhard Töpfer sei ebenfalls mein herzlicher Dank für die Prüfung meiner Arbeit und ihre fachliche Unterstützung ausgesprochen.

Dr. Lasse Klingbeil sei sehr für seine fortlaufende Unterstützung und Ratschläge über die Jahre am IGG, sowie die Korrekturlesung dieser Arbeit gedankt. Weiter möchte ich den Projektpartnern in NoViSys danken, insbesondere Markus Wieland, Dr. Katja Herzog und Dr. Anna Kicherer. Die gute Zusammenarbeit aller Beteiligten hat maßgeblich zum Erfolg des Projekts und dieser Arbeit geführt.

Ich danke allen aktiven und ehemaligen Mitarbeitern und Studenten am IGG für eine gute und lehrreiche Zeit. Insbesondere Dr. Stefan Paulus und Dr. Jan Dupuis sei für ihre fachliche Unterstützung und Freundschaft gedankt.

Ich danke nicht zuletzt meinen Eltern, Dr. Andreas und Christa Rose, für ihre liebevolle und fachliche Unterstützung über die Jahre als Student und Doktorand. Ihnen und meiner lieben Kristina, die mich immer, aber vor allem über die letzten Monate erduldet und umsorgte, kann ich nicht genug danken. Diese Arbeit ist ihnen und meinem verstorbenen Großvater, ÖbVI Winfried Rose, gewidmet.

**Von Wort zu Wort, von Werk zu Werk
- Hávamál**