

## LiDAR- und Lichtschattensensor-basierte Sensordatenfusion zur feldbasierten Phänotypisierung von Getreide

Iuliia Cherednyk<sup>1</sup>, Mario Jenz<sup>1</sup>, Kim Möller<sup>1</sup>, Dominik Nieberg<sup>1</sup> und Arno Ruckelshausen<sup>1</sup>

**Abstract:** In diesem Dokument werden komplementäre Sensorfusionsansätze zur feldbasierten Hochdurchsatz-Phänotypisierung von Triticale auf Grundlage bildgebender LiDAR- und Lichtschattensensor-Daten vorgestellt. Die Sensordaten wurden mit der Multisensor-Plattform „BreedVision“ gewonnen. Zur Bestimmung des parzellenbezogenen Parameters werden die Messdaten von LiDAR- und Lichtschattensensor fusioniert. Mit Methoden des überwachten maschinellen Lernens wird ein Biomasse-Vorhersagemodell erstellt. Die extrahierten Pflanzenmerkmale werden auf feldbasierte Daten von 1503 Versuchspartnern an zwei Versuchsstandorten für je zwei Wachstumsstadien trainiert. Das exponentielle Gauß'sche Prozessregressionsmodell wird mittels der Funktionen in MATLAB® entwickelt.

**Keywords:** LiDAR, Lichtschattensensor, Hochdurchsatz-Phänotypisierung, Sensordatenfusion, Machine Learning

### 1 Einleitung

Mobile Multisensor-Plattformen ermöglichen schnelle und zerstörungsfreie Aufnahmen von Sensordaten in Pflanzenkulturen unter Feldbedingungen und helfen den Züchtern bei der Beurteilung der Resistenz gegen Umweltbedingungen, Krankheiten und Schädlingen [Ra2015].

Die feldbasierte Phänotypisierung stellt große Anforderungen an die Qualität der Sensorrohdaten unter dem Einfluss von umweltbedingten Störgrößen, deren systematische („online“) Ablage und die („offline“) Interpretation großer Datenmengen zur Reduktion auf Pflanzenparameter [Ru14]. Die Auswertung basiert sowohl auf Algorithmen für einzelne Sensorrohdaten als auch auf der Multisensor-Datenfusion.

Die auf dem BreedVision-Konzept basierende Phänotypisierungsplattform (Abb. 1 A) beinhaltet zahlreiche optische Sensoren für die Vermessung der Pflanzen sowie Hilfssensorik zur Navigation und Positionsbestimmung [Bu13], [Mö16]. Die Sensorrohdaten der Feldversuche unter Einsatz der Multisensor-Plattform werden hochaufgelöst mit einem Orts- und Zeitstempel in einer Datenbank gespeichert. Die weitere Auswertung der Sensordaten sowie die Fusion unterschiedlicher Sensoren wird

durch die Verwendung der an der Hochschule Osnabrück entwickelten MATLAB® Applikation „DataAnalyzer“ ermöglicht. Diese Software beinhaltet mehrere Plugins zur Auswertung der Rohsensordaten [Mö16] und stellt eine Basis für die Datenfusion-Ansätze zur Berechnung von Pflanzenmerkmalen dar.

## 2 Material und Methoden

Die LiDAR-Sensoren liefern die Abstandsinformationen zwischen dem Messobjekt und dem Sensor und aufgrund dessen lassen sich die morphologischen Merkmale erfassen [Qi18]. Der Lichtschattensensor hat sich bereits als zuverlässiger Sensor zur Bestimmung von Höhe, Breite, Abstand und Bestandesdichte der Pflanzen etabliert [Bu13], [Ru14]. In dieser Arbeit wurde der 2D-LiDAR-Sensor LMS400 der Fa. SICK verwendet und der Lichtschattensensor RapidoScan RS-C-025-1760-ECT der Fa. dieEntwickler. In der Multisensor-Plattform „BreedVision“ ermittelt der LiDAR-Sensor die dreidimensionale Eindringtiefe von oben (Abb. 1 C), der Lichtschattensensor erfasst das zweidimensionale Pflanzenprofil seitlich innerhalb der Pflanzenparzelle (Abb. 1 B).

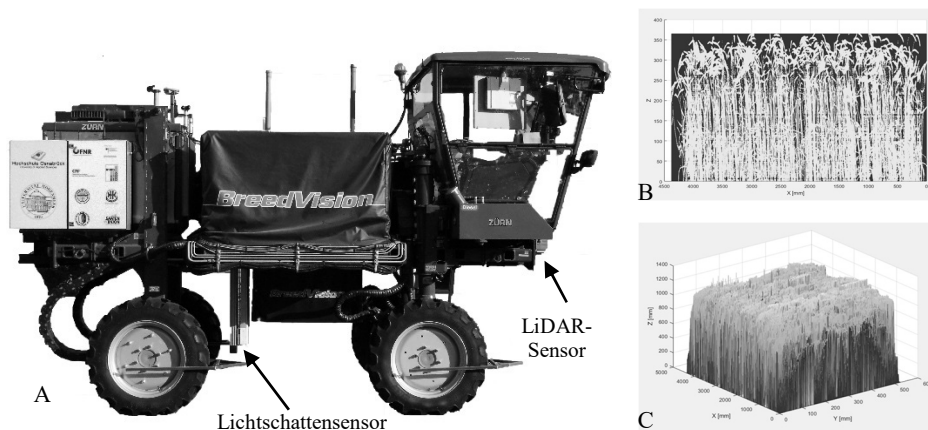


Abb. 1: A) Multisensor-Plattform „BreedVision“, B) Schattenbild von einer Parzelle aufgenommen mit Lichtschattensensor, C) Punktwolke einer Triticale-Parzelle, erfasst mit LiDAR

In diesem Kapitel wird die Berechnung der Pflanzenparameter anhand einzelner Sensordaten des 2D-LiDAR und Lichtschattensensors sowie deren Fusion vorgestellt. Als Quelle für Sensordatenfusion werden die Informationen sowohl von Rohmessdaten (Datenfusion) als auch von definierten Merkmalen von Objekten (Merkmalfusion) verwendet.

## 2.1 Bedeckungsgrad mit Lichtschattensensor

Der Bedeckungsgrad wurde als Prozentsatz der unterbrochenen Lichtschranken des Lichtschattensensors berechnet. Um eine mögliche Beeinträchtigung von Parzellen mit niedrigen Pflanzhöhen zu vermeiden, bestimmt der Algorithmus zunächst das Pflanzdach der Parzelle und berücksichtigt dann nur die Lichtschranken unterhalb der Oberkante der Pflanzen [Bu13].

## 2.2 Volumenbestimmung mit LiDAR –Sensor

Da der Laserscanner punktuell das Umgebungsprofil in zweidimensionalen Polarkoordinaten abtastet, werden die Messwerte ins kartesische Koordinatensystem umgerechnet. Eine Zusammensetzung der 2D-Sensordaten des LiDAR-Sensors und deren temporärer Positionen bildet die Grundlage, um die Entfernungs- und Remissionswerte von einer Versuchsparzelle dreidimensional darstellen zu können. Hierbei sind die Achsen mit den folgenden Koordinaten in Zusammenhang gebracht (Abb. 1 C): X-Koordinate – die Position des Laserscanners in Bewegungsrichtung der Phänotypisierungsplattform; Y-Koordinate – der Abstand zwischen dem Anfangs- und Endpunkt des Scans (bzw. einer Scan-Linie); Z-Koordinate – der Abstand vom Nullpunkt der Messung zur Oberkante des Objektes. Neben den gemessenen Entfernungswerten liefert der LiDAR-Sensor zu jedem Messpunkt auch Remissionswerte. Diese Messungen fügen der räumlichen Information des Sensors eine optische Erscheinungskomponente hinzu. Diese können verwendet werden, um innerhalb einer Parzelle die Pflanzenbestände vom Boden zu unterscheiden. Für die Erhöhung der Genauigkeit der Berechnungsalgorithmen über die gesamte Versuchsparzelle wird die binäre pixelbasierte Segmentierungsmethode nach dem Schwellwertverfahren angewendet. Der dreidimensionale Raum zwischen der Oberfläche der Vegetationsbedeckung und der Boden-Ebene einer Parzelle wird als parzellenbezogene Volumen bezeichnet. Das Volumen wird direkt aus jedem isolierten Punkt in der Punktwolke berechnet [ZM05]. Dabei wird der 3D-Raum in Quader mit der Dimension  $(dh_{x,y}, dx, dy)$  eingeteilt. Das Volumen wird durch das Aufsummieren aller beibehaltenen Volumenelemente berechnet:

$$V = \int_V dV = \iint h_{x,y} dx dy = \sum \sum h_{i,j} \Delta x_i \cdot \Delta y_j . \quad (1)$$

$h_{i,j}$  entspricht den positionsabhängigen Höhenwerten des Pflanzenbestandes. Durch die Positionsänderung über der x -Achse  $\Delta x$  und die Positionsänderung über die y-Achse  $\Delta y$  wird ein Flächenelement bestimmt.

## 2.3 Sensorfusion LiDAR – und Lichtschattensensor auf Datenebene

Bei der Fusion auf Datenebene werden die Sensorrohdaten des LiDAR- und Lichtschattensensors direkt miteinander kombiniert. Die beiden Sensoren arbeiten mit unterschiedlichen Abtastfrequenzen. Um eine Datenfusion zu ermöglichen, müssen die

Messdaten mit einer gemeinsamen Ortsauflösung (die Position auf der X-Achse) referenziert werden. Bei der Berechnung des Volumens einer Parzelle ist zu beachten, dass der Lichtschattensensor die seitliche Flächenbelegung in der Mitte einer Parzelle (zwei Saatzeilen) erfasst. Es wird angenommen, dass die anderen Saatzeilen in dieser Parzelle die gleiche Dichte haben. Die nichtunterbrochenen Lichtstrahlen im Schattenbild zeigen die Lücke im Flächeninhalt der begrenzten Flächen zwischen den Boden- und Höhenpunkten einer Scan-Linie des LiDAR-Sensors in der x-Position:

$$F_x = \sum (\Delta F_x \cup N_{lg\ x}), \quad (2)$$

mit den Teilflächen an einem Standpunkt der X-Achse  $\Delta F_x$  und einem Vektor des Schattenbildes in x-Position  $N_{lg\ x}$ . Das Volumen einer Parzelle wird als Summe der Multiplikation aller Teilflächen und der allgemeinen Auflösung der X-Achse berechnet:

$$V = \sum \left( \begin{bmatrix} F_{x1} \\ \dots \\ F_{xi} \end{bmatrix} \cdot \Delta x \right). \quad (3)$$

#### 2.4 Sensorfusion LiDAR – und Lichtschattensensor auf Merkmalebene

Die Fusion auf Merkmalebene folgt nach einer Vorbearbeitung, wodurch die Datenmenge auf Merkmale reduziert wird. Hierbei werden die beiden pflanzlichen Merkmale Höhe und Dichte in den jeweiligen Positionspunkten der X-Achse zusammengelegt. Hier wird die positionsreferenzierte Dichte als unterbrochene Zellen zur gesamten Anzahl der unter den höchsten Werten liegenden Pixeln des Schattenbildes bestimmt. Das Volumen wird als Summe der Volumenelemente, die mit der positionsreferenzierten Dichte von Lichtschattensensor gewichtet werden, berechnet:

$$V_{veg} = \sum \sum \left[ \begin{bmatrix} h_{11} & \dots & h_{1j} \\ \dots & \dots & \dots \\ h_{i1} & \dots & h_{ij} \end{bmatrix} \circ \begin{bmatrix} \Delta y_{11} & \dots & \Delta y_{1j} \\ \dots & \dots & \dots \\ \Delta y_{i1} & \dots & \Delta y_{ij} \end{bmatrix} \circ \begin{bmatrix} \rho_1^* \\ \dots \\ \rho_i^* \end{bmatrix} \right] \Delta x. \quad (4)$$

#### 2.5 Biomassevorhersage mittels Machine Learning

Die zerstörungsfreie Bestimmung der Biomasse von einer Triticale-Parzelle ist eine komplexe Aufgabe, die viele präzise Sensordaten benötigt. Da keine definierten Abhängigkeiten dafür vorhanden sind, werden die Methoden des überwachten Lernens für die Erstellung des Frischbiomasse-Vorhersagemodells verwendet. Die konventionell erfassten Biomassedaten lagen als Referenzwerte zum Trainieren des Modells vor. Als Training-Datensatz werden die aus den feldbasierten Sensorrohdaten extrahierten Pflanzenmerkmale Volumen, Bedeckungsgrad und Wuchshöhe verwendet. Das Vorhersagemodell wird mittels der Gauß'schen Prozessregression mit der exponentiellen Kernel-Funktion `fitrgp()` von MATLAB® erstellt.

### 3 Ergebnisse und Diskussion

Als Anwendungsbeispiel der Sensordatenfusions-Algorithmen wird die Biomasse von Triticale-Versuchspartellen betrachtet. Hierzu liegen jeweils invasiv konventionell erfasste Frischmassereferenzdaten zum Vergleich und zur Kalibrierung vor.

Die Auswertung erfolgte auf vier Datensätze: zwei Wachstumsstadien an zwei unterschiedlichen Orten (Tab. 1). Die Datensätze umfassen insgesamt 1503 Feldversuchspartellen. Es wurden hohe Determinationskoeffizienten für den Vergleich zwischen dem parzellenbezogenen Volumen und Frischmassereferenzdaten erreicht: Die besten Ergebnisse wurden für die LiDAR- und Lichtschattensensor-basierende Datenfusion auf Merkmalsebene erzielt.

Sensor	Parameter	Stuttgart-Hohenheim		Moosburg	
		25.05.2018	20.06.2018	4.06.18	14.06.18
LS	Bedeckungsgrad	0,7105	0,7975	0,7327	0,7742
LiDAR	Volumen	0,7029	0,8624	0,8724	0,7382
LiDAR und LS	Volumen (Daten Fusion)	0,6754	0,8702	0,8782	0,7555
LiDAR und LS	Volumen (Merkmal Fusion)	0,7862	0,8979	0,9033	0,8812

Tab. 1: Die Determinationskoeffizienten der Regression zwischen dem Volumen und der Frischmasse für zwei unterschiedliche Wachstumsstadien an zwei Versuchsorten

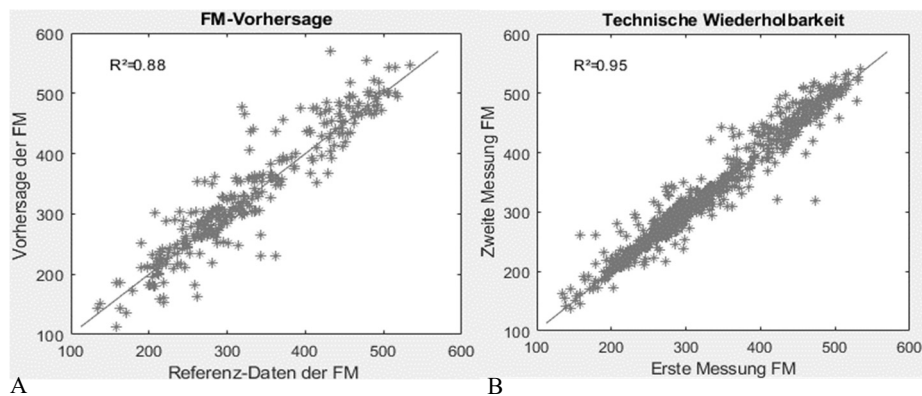


Abb. 2: A) Korrelation zwischen vorhergesagten Daten mit Machine Learning und Referenz-Daten; B) Technische Wiederholbarkeit der vorhergesagten frischen Biomasse

Jeder Parzelle wurde mehrfach vermessen, um die technische Wiederholbarkeit der Messsensoren selbst und der Algorithmen für die Bestimmung der Pflanzenmerkmale zu validieren. Das Bestimmtheitsmaß der Wiederholung der einzelnen Parameter liegt

zwischen 0,84 und 0,95. Als Training-Datensatz für das Biomasse-Vorhersage-Modell werden 80 % der extrahierten Parameter der Feldversuchspartellen verwendet. Die Restlichen 20 % werden zum Testen und Berechnung der Genauigkeit gebraucht. Das Bestimmtheitsmaß zwischen den vorhergesagten und den Referenz-Daten beträgt 0.88 (Abb. 2A). Die Genauigkeit der technischen Wiederholbarkeit liegt bei 0.95 (Abb. 2B).

Durch die Verwendung zusätzlicher Pflanzenparameter (z. B. Wassergehalt) oder eine Vergrößerung des Training-Datensatzes könnte die Genauigkeit des Vorhersage-Modells noch weiter erhöht werden.

## 4 Schlussfolgerungen

Es wurden LiDAR- und Lichtschattensensor-basierte Sensordatenfusion-Ansätze anhand von Merkmalen und Sensorrohdaten betrachtet. Auch wenn Sensoren allein betrachtet eine niedrigere Korrelation aufweisen, kann diese durch Sensordatenfusion deutlich verbessert werden. Das berechnete parzellenbezogene Volumen auf Merkmal-Ebene erzielt eine bessere Korrelation mit frischer Biomasse als auf Datenebene. Die vorgestellte Methodik bietet das Potenzial für Anwendungen mit weiteren Sensoren und Merkmalen.

**Danksagung:** Das Forschungsprojekt SENSELGO wird vom Bundesministerium für Ernährung und Landwirtschaft, der Fachagentur für nachwachsende Rohstoffe e.V. und der Gemeinschaft zur Förderung von Pflanzeninnovationen e.V. gefördert.

### Literaturverzeichnis

- [Bu13] Busemeyer, L. et al.: BreedVision — A Multi-Sensor Platform for Non-Destructive Field-Based Phenotyping. In *Plant Breeding. Sensors* 13(3), Basel (Switzerland), S. 2830-2847, 2013
- [Mö16] Möller, K., et. al.: Feldtaugliche Multisensorplattform für High-Throughput Getreidephänotypisierung – Aufbau und Datenhandling. In: Ruckelshausen, A., Meyer-Aurich, A., Rath, T., Recke, G. & Theuvsen, B. (Hrsg.): *Informatik in der Land-, Forst- und Ernährungswirtschaft 2016*. Bonn: GIL e.V., S. 137-140, 2016
- [Qi18] Qiu, R., Wei, S., Zhang, M., Li, H., Sun, H., Liu, G., & Li, M.: Sensors for measuring plant phenotyping: A review. *International Journal of Agricultural and Biological Engineering*, 11(2), S. 1-17, 2018.
- [Ra15] Rahaman, M., Chen, D., Gillani, Z., Klukas, C., & Chen, M.: Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Frontiers in plant science* vol. 6 619. 2015
- [Ru14] Ruckelshausen, A.: Neue Sensorentwicklungen – ein technischer Blick auf Pflanzen, Sensoren und Daten. In *Journal für Kulturpflanzen*, 66(2), Stuttgart, S. 73-79, 2014.
- [ZM05] Zhang, X., & Morris, J.: Volume measurement using a laser scanner. CITR, The University of Auckland, New Zealand, 2005