Funktionsweise und Vergleich von Methoden zur Generierung von Punktwolken in einer Versandstation

Doran Holger Scholl Fakultät Informatik Hochschule Furtwangen University Furtwangen, Deutschland scholldo@hsfurtwangen.de Dominik Sandro Schwer Fakultät Informatik Hochschule Furtwangen University Furtwangen, Deutschland schwerdo@hsfurtwangen.de

Zusammenfassung-Diese Ausarbeitung präsentiert und bewertet verschiedene Methoden zur Erzeugung von Punktwolken anhand eines Anwendungsfalls aus der Industrie. Als zu erkennendes Objekt wurde ein Paket innerhalb einer Versandstation definiert. Untersucht werden eine Time-of-Flight (TOF), eine Tiefen-Stereo, und eine aus zwei Low-Budget Webcams bestehende, selbst gebaute Stereokamera. Hierfür wurden Punktwolken aus den Aufnahmen der Kamerasysteme, welche denselben Szenen und Modellparametern entsprechen, erzeugt. Die Evaluierung beinhaltet einen Bewertungskatalog sowie ein Bewertungsvorgehen nach 5-Sterne-Rating-Schema. Die enthaltenen Kriterien werden in Form eines Fünf-Sterne-Rating-Schemas beurteilt. Die Ergebnisse der Evaluation zeigen, dass der TOF-Ansatz für den domänenspezifischen Anwendungsfall das beste Gesamtergebnis in Höhe von 4,7 Sternen liefert. Sämtliche technischen Anforderungen wurden zur vollen Zufriedenheit umgesetzt, nur der Kostenaspekt erwies sich als nachteilig. Der Einsatz eines Tiefen-Stereo-Systems folgt mit 3,7 Sternen und bestätigt somit eine Applikabilität für den Aufwendungsfall. Allerdings belegen die Untersuchen eine Vulnerabilität des Systems durch Störfakten wie reflektierende Oberflächen. Der Selfmade-Ansatz erwies sich mit einem Gesamtergebnis von 2,1 Sternen als unbrauchbar, da die technischen Anforderungen nicht erfüllt wurden. Bedingt durch die geringen Kosten des Systems kann dieser Ansatz zu Lehr- und Forschungszwecken in Betracht gezogen werden. Die Untersuchung hat ergeben, dass sich der TOF-Ansatz und der Stereo-Ansatz für einen industriellen Einsatz für die Formerkennung von Paketen in einer Versandstation eignen.

Keywords—stereo camera, time-of-flight, depth map, pointcloud

I. MOTIVATION

Die voranschreitende Entwicklung der Industrie 4.0 ist ein Treiber für die zunehmende Bedeutung von autonomen Robotern [1, p. 23]. Damit verbunden steigen die Anforderungen an Kamerasysteme, da die erzeugten Aufnahmen mithilfe von Objekterkennungsalgorithmen das Arbeiten mit autonomen Robotern ermöglichen. Im Anwendungsfall einer industriellen Versandstation müssen Roboter in der Lage sein, unterschiedliche Objekte wie Pakete in unterschiedlichen Positionen zu erkennen, um mit diesen in geeigneter Form zu interagieren [2, p. 1] [3, p. 1]. Um dies zu ermöglichen, werden Kamerasysteme eingesetzt, welche in der Lage sind Bilder der Szene zu erfassen, zu analysieren und die Informationen in geeigneter Form an den Roboter weiterzuleiten. Die Herausforderung besteht allerdings darin, eine Richard Hempel Fakultät Informatik Hochschule Furtwangen University Furtwangen, Deutschland hempelri@hsfurtwangen.de Anton Petrov Fakultät Informatik Hochschule Furtwangen University Furtwangen, Deutschland petrovan@hsfurtwangen.de

2D-Aufnahme in ein für den Roboter geeignetes 3D-Format darzustellen [4, pp. 236-237]. In diesem Kontext existiert eine Vielzahl von Verfahren und Methoden, um eine 2D-Aufnahme in ein für den Roboter verwendbares 3D-Bild umzuwandeln [5, pp. 1252-1527]. Eine Grundlage für dieses Verfahren ist das Erzeugen einer Punktwolke die eine unsortierte Sammlung von 3D-Koordinaten oder Ortsvektoren im dreidimensionalen Raum darstellt [6, pp. 240-241].

Anwendungsfall dient ein Paket in einer Als Versandstation, welches in Form und Größe erfasst werden soll. Mit dem Ziel, eine Bewertung der Kamerasysteme unter Berücksichtigung des Anwendungsfalles durchzuführen, wurden je Kamerasystem Bilder mit identischen Modellparametern, d. h. Positionen, Abständen und Winkeln aufgenommen. Innerhalb dieser Ausarbeitung wurde erforscht, welches Kamerasystem dem Anwendungsfall entsprechend als besonders geeignet gilt. Dafür werden zu Beginn als Grundlage die Systeme erläutert. Darauffolgend wird ein Bewertungskatalog vorgestellt. Anhand dieses Bewertungskatalogs werden die vorgestellten Systeme bewertet. Das Fazit markiert die Auswertung der Systeme.

II. GRUNDLAGEN

Im Nachfolgenden wird die Funktionsweise der ausgewählten Kamerasysteme zur Erzeugung einer Punktwolke beschrieben.

A. Funktionsweise zur Erzeugung von Tiefeninformationen durch Triangulation

Durch den Einsatz einer Stereokamera lassen sich unter Verwendung von intrinsischen Kameraparametern wie dem Öffnungswinkel, die Brennweite und der Chipgröße, die Pixelentfernung im dreidimensionalen Raum in Zentimeter umrechnen. Die Grundlage bildet die Epipolargeometrie, die beiden Kameras und deren Bildebenen beschreibt. Daraus ergibt sich, dass korrespondierende Punkte auf den jeweiligen Epipolarlinien liegen müssen. Somit ist die Triangulation ist die Berechnung des Schnittpunkts der beiden Geraden durch die korrespondierenden Pixel. In dieser Ausarbeitung wird ein Verfahren in vier Schritten beschrieben [7, p. 755].

1) Rektifizierung: Der Grundgedanke des Verfahrens zur Generierung einer Punktwolke mithilfe des Einsatzes einer Stereokamera basiert auf der Logik zweier Kameraaufnahmen. Das im Raum stehende zu ermittelnde Objekt wird von zwei Kameralinsen erfasst. Bei der erzeugten Auffnahme müssen im Anschluss die jeweiligen korrespondierenden Pixel (Pixel, die denselben Punkt im dreidimensionalen Raum abbilden) ermittelt werden. Die Abbildung 1 veranschaulicht dabei die fundamentale Funktionsweise. Die Punkte C₁ und C₂ beschreiben die bekannten Positionen der Kamera, X₀ und X₁ sind die korrespondierenden Punkte und P ist der Punkt im dreidimensionalen Raum, den es zu berechnen gilt, da dieser Punkt die physischen Koordinaten aus der Abbildung in X₀ und X₁ besitzt. Damit dies möglich ist, muss die Position, Ausrichtung, Brennweite und der Kamerawinkel beider Kameras bekannt sein (R,t). Das Verfahren funktioniert wie folgt [7, pp. 755-756]:

- ziehe eine Gerade unbestimmter Länge von C₀ durch X₀
- ziehe eine Gerade unbestimmter Länge von C₁ durch X₁
- der Schnittpunkt dieser beiden Geraden ist der Punkt P, seine Position bestimmt den Punkt im dreidimensionalen Raum, der durch die korrespondierenden Pixel dargestellt wird
- die Abbildung der Linie C₀ und P∞ in der Ebene eins und die Abbildung der Linie C₁ P in der Ebene Null bilden korrespondierende Epipolarlinien



Abb. 1. Funktionsweise einer Stereokamera (Rohdarstellung) [7, p. 755]

Abbildung 2 zeigt das Rektifizierungsverfahren. Ziel dieses Verfahrens ist es, alle korrespondierenden Pixel auf einer Höhe zu ermitteln, dabei müssen die Epipolarlinien den horizontalen Scanlinien entsprechen. Durch den Einsatz der Epipolarlinien müssen nun nur noch alle Punkte einer Linie und nicht alle Punkte auf jedem Bild verglichen werden. Dies wird durch das Verzerren der Bilder ermöglicht, dabei sollen die horizontalen Scanlinien den Epipolarlinien entsprechen. Die Epipolarlinien sind Linien, auf denen korrespondierende Punkte existieren können. Die horizontalen Scanlinien [7, p. 756].



Abb. 2. Rektifizierung -Verfahren [7, p. 756].

Um den Tiefenwert im dreidimensionalen Raum zu berechnen wird die Formel 1 verwendet.

$$Z = f \frac{B}{d} \tag{1}$$

Der Wert Z ist die Tiefe im dreidimensionalen Raum und stellt somit die Entfernung der Kamera von jedem Pixel des zu erkennenden Objekts im Bild dar (Depth-Map), die Variable f beschreibt die Brennweite der Kamera und ist somit hardwarespezifisch gegeben, die Baseline B ist der relative Abstand der Kameras und d stellt die Distanz der Pixel dar [7, p. 755]. Die Formel 2 ist die Grundlage für die Berechnung der Entfernung von den Pixel zur Kamera, dies wird für die Generierung der Punktwolke benötigt. Die umgestellte Formel lautet wie folgt [7, p. 755]:

$$d = f \frac{B}{Z} \tag{2}$$

Die Formel 3 beschreibt die Annahme aus der Rektifizierung zur Position und Form der Epipolarlinien [7, p. 755]. Das Pixel links sowie das Pixel rechts liegen auf einer Höhe (gleiche y-Koordinate) aber nicht auf derselben Breite (unterschiedliche x-Koordinaten).

$$x' = x + d(x, y), y' = y$$
 (3)

2) Bestimmung von korrespondierenden Pixel: Nachdem das Rektifizierungsverfahren angewandt wurde, muss werden. identifiziert welche Pixel miteinander korrespondieren, da nur ermittelt werden konnte, welche Linien potenziell miteinander korrespondieren. Hierfür werden die bereits erstellen Epipolarlinien genutzt. Ein Verfahren, um die exakten korrespondierenden Pixel zu identifizieren, ist die Fotokonsistenz, bei dem jedes Pixel aus der linken Linie mit jedem Pixel aus der rechten Linie verglichen wird. Durch die Similarity Measure findet eine Ähnlichkeitsmessung statt, dabei werden jene Pixel innerhalb Epipolarlinie gemappt, welche der den höchsten Ähnlichkeitsfaktor besitzen. Im einfachsten Fall ist dies die absolute Differenz der Farbwerte in einem Fenster um das Pixel herum. Jene Pixel werden nun als korrespondierende Pixel behandelt [7, pp. 755-758]. Es lässt sich zusammenfassen, dass dank der Rektifizierung nur in einer Zeile gesucht werden muss. Darin werden die Pixelwerte aus einem Fenster um das zu prüfende Pixel herum verglichen. Das ähnlichste Pixel wird als korrespondierend behandelt.

3) Generierung einer Depth-Map: Eine Depth-Map ist eine 2D-Abbildung einer Szene, bei der die Entfernung jedes Punkts zur Kamera in einem Bild codiert ist. Die Abbildung 3 zeigt die Darstellung einer Depth-Map. Dabei handelt es sich um eine Tiefenkarte, welche Informationen über die Entfernung der Oberfläche eines Objektes aus Sicht der Kamera enthält. Dabei enthält jedes Pixel keinen RGB-Farbwert (a) sondern die Entfernung zum Objekt. Je heller das Pixel, desto näher ist das Objekt in Relation zur Kamera (b). Um aus dem erzeugten Bild der Stereokamera inkl. Epipolarlinien eine Depth-Map zu erzeugen, kann die Formel 1 zur Berechnung der Tiefe im dreidimensionalen Raum aus dem genutzt werden. Nachdem jedes korrespondierende Pixel in die Formel eingesetzt wurde, lässt sich die Depth-Map erzeugen [7, pp. 757-758].



Abb. 3. Darstellung einer Depth-Map [7, p. 757].

4) Erstellung einer Punktwolke: Eine Punktwolke ist eine Sammlung von Punkten im 3D-Raum, die verwendet werden kann, um die Geometrie von Objekten oder Umgebungen zu beschreiben. Um eine Punktwolke aus einer Depth-Map zu müssen zunächst die intrinsischen erzeugen, Kameraparameter bekannt sein, d.h. die Eigenschaften der Kamera, wie zum Beispiel die Brennweite und der Sensorabstand. Die Depth-Map wird dann analysiert und für jedes Pixel ein 3D-Punkt erzeugt, indem die Entfernung zur Kamera und die 2D-Position des Pixels in der Depth-Map verwendet werden. Um die 3D-Punkte zu erzeugen, werden in der Regel Triangulationsverfahren verwendet, die es ermöglichen, die 3D-Position eines Punkts anhand seiner 2D-Position und der bekannten Kamera-Kalibrierung zu berechnen. Einmal erzeugt, kann die Punktwolke verwendet werden, um die Geometrie der Szene zu analysieren oder um die Szene in 3D zu visualisieren. Sie kann auch verwendet werden, um dreidimensionale Modelle von Objekten oder Umgebungen zu erstellen, indem die Punkte in der Punktewolke analysiert und miteinander verbunden werden [16].

B. Funktionsweise zur Erzeugung von Tiefeninformation durch Lichtlaufzeitmessung

Bei einer Time-of-Flight-Kamera handelt es sich um ein Tiefenkamerasystem, das auf Basis der Lichtlaufzeit die Distanz der Objekte zur Kamera erfasst. Unter Verwendung eines TOF-Sensor-Chip ist es möglich, für jedes Pixel einen Distanzwert zu ermitteln. Durch Verwendung von intrinsischen Kameraparametern kann anschließend eine Punktwolke erzeugt werden.

In diesem Abschnitt wird die Generierung einer Punktwolke mittels einer TOF-Kamera erörtert.

1) TOF-Prinzip: Um die relative Distanz zwischen der Kameralinse und dem zu erkennenden Objekt festzustellen, nutzt dieser Ansatz die Ausbreitungsgeschwindigkeit des Lichts, um die Distanz zu einem Objekt zu berechnen. Dabei wird die Lichtlaufzeit vom Sender zum Objekt und zurück zum Empfänger gemessen. Durch die TOF-Kamera werden Lichtaussendungen auf das im Raum stehende Objekt ausgesendet, welches es wieder reflektiert und von der Kamera erneut aufgenommen wird. Unter Betrachtung des Lichtlaufzeitverfahrens kann die Entfernung zu jedem Pixel im dreidimensionalen Raum berechnet werden [8, p. 808]. Dieses Verfahren lässt sich durch den Einsatz zweier Lichtlaufzeitverfahren grundlegende umsetzen. Der Unterschied der Methoden wird in der Funktionsweise der Modulation des Lichtsignals ersichtlich.

Bei dem Pulse-Based (PB) Modulationsverfahren sendet die Beleuchtungseinheit des Kamerasystems in kurzen Abständen einzelne Lichtimpulse auf das Zielobjekt. Anschließend nimmt der Empfänger die reflektierten Lichtimpulse des Zielobjekts auf [9, p. 4]. Bei dem Continuous-Wave (CW) Modulationsverfahren hingegen sendet die Beleuchtungseinheit des Kamerasystems einen durchgehenden Wellenimpuls, welcher durch die Frequenzmodulation eine variable Helligkeitsform besitzt, auf das Zielobjekt. Durch die Reflektion des Zielgegenstandes sind die versendeten und empfangenen modellierten Lichtimpulse phasenverschoben. In zwei Schritten berechnet der Tiefensensor die Entfernung zu dem Zielobjekt aus der Messung der Phasenlage Phasendifferenz, also der reflektierenden helligkeitsmodulierten Lichtquelle des Bildes [10, pp. 21-25]. Für die Berechnung der Phasenverschiebung φ muss zunächst die Kreuzkorrelation zwischen dem entsendetem und dem empfangendem Lichtimpuls betrachtet werden [10, pp. 21-25].

$$\varphi = \operatorname{atan}(\frac{Q_3 - Q_4}{Q_1 - Q_2}) \tag{4}$$

Im zweiten Schritt wird die Distanz d für jedes Pixel berechnet. Die Formel 5 zeigt den Quotienten d. Der Dividend, die Lichtgeschwindigkeit c, wird durch den Divisor die Modulationsfrequenz $4\pi f$ geteilt. Hierfür gelten jeweils $2\pi f$ für den Hin- und Rückweg des modulierten Lichts. Als Ergebnis erhält das System die relative Distanz ohne Berücksichtigung der Phasenverschieben. Um die φ zu berücksichtigen, wird der erste Quotient mit der φ multipliziert. Somit ist die Distanz proportional zur Phasenverschiebung [9, p. 23].

$$d = \left(\frac{c}{4_{\pi f}} \, \varphi\right) \tag{5}$$

2) Umsetzung durch das Kamerasystem: Die Hardwareschicht eines TOF-Kamerasystems besteht aus einem Emitter, einer Optik, einem TOF-Sensor und einem Control Panel, siehe Abb. 4. Der Emitter besteht aus einem Infrarotlicht Transmitter, welcher als Festkörperlaser oder einer Leuchtdiode Platz findet. Der Emitter enthält eine Modulationsquelle, um die Frequenz des Transmitters zu steuern. Zudem muss der Emitter in der Lage sein, hohe Frequenzen zu senden, um mit moduliertem Licht die Szene zu scannen [9, p. 19]. Die Optik ist für das Empfangen des reflektierten Infrarotlichts vom Zielobjekt zuständig. Dies wird durch Linsensysteme und einem Intensitätdetektor, der den reflektierten Lichtimpuls misst, ermöglicht. Der Tiefensensor besteht aus mehreren Photodetektoren, welche reagieren, wenn Lichtsignale nach der Reflektion erneut empfangen werden. Die Entfernungsmessung wird anhand der Zeit berechnet, welche nach Austritt des Emitter Signals und dem Empfang der Reflektion vergeht. Da die Lichtgeschwindigkeit konstant ist, lässt sich die Entfernung aus der Dauer des Weges messen. Das Control Panel ist das zentrale Hilfsmittel für die Systemsteuerung, wodurch der Ablauf der Befehlsverarbeitungen durchgeführt wird [8, p. 20].



Abb. 4. Funktionsweise und Hardwareschicht eines TOF-Kamerasystems [Anlehnung an 8, p. 20]

3) Generierung einer Depth-Map: Um ein Tiefenbild mittels TOF-System zu generieren, muss der Emitter Licht in die ganze Szene entsenden und parallel alle Pixel gleichzeitig verarbeiten. Bei dem PD-System erfolgt dies über einzelne Lichtimpulse. Dabei berechnet es die Tiefeninformationen der Pixel anhand der Zeit, die zwischen Versenden und Empfangen der einzelnen Lichtimpulsen vergeht. Bei dem CW-System erfolgt dies über den Einsatz von durchgehenden Wellenimpulsen, dabei berechnet dieses System die Tiefeninformationen anhand der Phasenlage durch die reflektierenden helligkeitsmodulierten Lichtquellen. Mit den gesammelten Tiefeninformationen der Pixel und den bereits vorhandenen intrinsischen Kameraparameter wie Brennweite kann nun eine Depth-Map aus den einzelnen Pixel generiert werden. Sofern dies als Grundlage gefestigt ist, kann unter der bererits vorgestellten Methode eine Punktwolke generiert werden.

C. Funktionsweise einer Webcam basierten Stereokamera

Das Triangulationsverfahren bildet die Grundlage zur Generierung einer Punktwolke mithilfe einer Webcam basierten und selbst gebauten Stereokamera. Dafür werden zwei Webcams durch ein USB-Kabel miteinander verbunden.

Um die Qualität dieses Systems sicherzustellen, müssen beide Webcams so kalibriert werden, dass sie die Beziehung zwischen den 3D-Koordinaten der betrachteten Szene und den zugehörigen 2D-Bildkoordinaten (Pixel) möglichst genau beschreiben. Die Schlüsselkomponenten des Verfahrens basieren weitgehend auf das in [10, pp.1489-1490] vorgestellte Verfahren.

Kalibrierungsprozess: Das Konzept der Kamerakalibrierung unterscheidet sich in extrinsisch und intrinsisch. Die extrinsische Kamerakalibrierung verantwortet das Finden der Position und der Orientierung der Kamera in einem Bezugsraum. Die intrinsische Kamerakalibrierung verantwortet das Finden der Abbildungsvorschrift der Kamera von 3D auf 2D, wie den Öffnungswinkel und die Brennweite. Die Kalibrierung kann durch das Bestimmen eines Punktes, welcher jeweilig in den verschiedenen 2D Bildern zu erkennen ist, durchgeführt werden. Der Punkt übermittelt dabei genug Informationen, um die Projektion dessen auf der Bildfläche abzustimmen. Hierbei sorgt die Projektionsabgleichung dafür, dass die intrinsischen Parameter der jeweiligen Kameras wie der Brennpunkt abgestimmt werden [11, pp. 15-34]. In Bezug auf die Kameralinsen Kalibrierung bedeutet das, dass beide Linsen

dasselbe Objekt an einer anderen Stelle sehen und schaffen somit die Grundlage für die Triangulation.

III. Methodik

Dieses Kapitel stellt das Forschungsvorgehen und die Evaluation der Ergebnisse vor.

A. Vorgehensmethodik

In diesem Abschnitt wird das Forschungsvorgehen erläutert, um eine reproduzier- und nachvollziehbare Studie durchzuführen. Im Folgenden werden die Kamerasysteme vorgestellt, mit denen die Aufnahmen erzeugt wurden.

Tiefen Stereokamera: bei der verwendeten Stereokamera handelt es sich um eine intel RealSense[™] d435. Sie besitzt eine Auflösung von 2MP und kann Tiefenerkennungen bis zu zehn Metern durchführen.

2) TOF-Kamera: bei der verwendeten TOF-Kamera handelt es sich um eine Azure Kinect DK des Herstellers Microsoft welche das CW-Verfahren zur Tiefenmessung anwendet. Durch den integrierten 3,5 μ m Megapixel TOF-Sensor ist sie in der Lage kleinste Pixel mit einer Genauigkeit im Bereich von 3,5 mm zu erfassen.

3) Webcambasierte Stereokamera: dabei handelt es sich um zwei Logitech c270 Webcams des Herstellers Logitech. Sie besitzen jeweils 0.9 Megapixel und eine Videoauflösung von 1280 x 720 Pixeln. Die Kalibrierung der Webcams wurde mit Hilfe von OpenCV 4.6 und der Bibliotek Stereovision 1.0.4 umgesetzt. Das dafür verwendete Schachbrettmuster besitzt 10x7 Quadrate von einer je Größe 1,8 cm.

Die Zielobjekte für den motivationsgetriebenen Anwendungsfall werden in der Tabelle 1 präsentiert.

 Tabelle I.
 BESCHREIBUNG DER ZIELOBJEKTE (PAKETE)

Name (ID)	Beschreibung	
P1	Maße: 23x12x5,5 cm	
P2	Maße: 23x15x8 cm	
Р3	Maße: 31,5x14,4x9,5 cm	

Die Darstellung des Aufbaus ist der Abbildung 5 zu entnehmen. Für alle Zielobjekte galten dieselben Bedingungen zur Erstellung der Aufnahmen. Es wurde darauf geachtet, dass die Lichtverhältnisse, der Untergrund sowie der Hintergrund des Aufbaus unverändert blieben. Um möglichst präzise Aufnahmen der Zielobjekte mit den verschiedenen Kamerasystemen zu ermöglichen, wurde auf der Oberfläche des Tisches eine Markierung für die jeweiligen Entfernungen und Positionen gesetzt. Es wurde jeweils aus den Entfernungen nah (40 cm), mittel (55 cm) und lang (70 cm) getestet. Ein 60° Winkel aus der parallelen Stellung zur Kamera ermöglichte das Einsehen von drei Flächen des Pakets. Ebenso wurde die Position der Kameralinsen auf der Tischoberfläche markiert. Die Höhe der Linse entsprach 17 cm und wurde durch die Nutzung eines Statives gewährleistet. Zudem wurde eine Reihenfolge zur Fotoerstellung definiert, da ein Paket nicht alle Zustände des Bewertungskatalogs darstellen konnte.



Abb. 5. Darstellung des Aufbaus

Die Erzeugung der Punktwolken basieren bei dem TOFund dem Stereoansatz über eine mitgelieferte Software. Parameter wurden Hierbei die angepasst. Die Schlüsselkomponenten des Verfahrens zur Kalibrierung des Selfmade-Ansatzes basieren weitgehend auf dem in [13, pp. 10-13] vorgestellten Vorgehen. Zunächst wurden die Webcams mit Klebeband auf einer Vorrichtung auf selber Höhe und selben Blickwinkel fixiert. Anschließend wurden die Webcams angeschlossen und deren Index definiert. Dadurch konnten mithilfe eines selbst geschriebenen Programms gleichzeitig Bildaufnahmen aufgenommen werden. Abbildung 6 zeigt den selbsterstellten werden. Abbildung 6 Kalibrierungsprozess mithilfe eines 9x13 Schachbretts zur Linsenkalibrierung [10, pp.1489-1490]. Nachdem die

Webcams zu einer Stereokamera kalibriert wurden, konnte mithilfe von [12, pp. 3-6] aus den Aufnahmen eine Punktwolke generiert werden. Die Betrachtung der Punktwolken erfolgt unter Verwendung des Softwaretools MeshLab.



Abb. 6. Kalibrierungsprozess für die Selfmadekamera

B. Evaluationsmethodik

In diesem Abschnitt wird die Methodik zur Bewertung der Kamerasysteme erläutert. Der in der Tabelle II definierte Bewertungskatalog ist das Ergebnis von intensiven Spezialisten-Interviews und Recherche. Der Katalog legt durch die Kriterien die Grundlage für die Bewertung der drei Kamerasysteme. Die Tabelle III definiert die Metriken für die Vergabe der Sterne. Dabei steht die Abkürzung RW für Referenzwert und AW für Abweichung.

Fabelle II. ERKLÄRUNG DER BEWERTUNGSKRITERIE	N
--	---

Nr.	Kriterium	Beschreibung	
1	Anzahl der Punkte	Aus wie vielen Punkten besteht die Punktwolke?	
2	Wiederholgenauigkeit	Führen wiederholte Messungen zu demselben Ergebnis?	
3	Relative Messgenauigkeit	essgenauigkeit Entspricht die virtuelle Messung der relativen Messgröße?	
4	Fähigkeit der Anpassung	Ist das System gegenüber Störfaktoren resilient?	
5	5 Berücksichtigung von Warm- und Kaltlaufphasen Beeinflusst die Betriebszeit die Qualität der Ergebnisse?		
6	Kosten bzw. Preis	Betrachtung der Anschaffungs- und Betriebskosten.	

Tabelle III. METRIKEN FÜR DAS BEWERTUNGSVERFAHREN

Nr.	*	* *	* * *	* * * *	* * * * *
1	0% - 39% des RW	40% - 59% des RW	60% - 79% des RW	80% - 100% des RW	Höchste Anzahl der Punkte
2	Mehr als 4 mm AW	3,1 mm - 4 mm AW	2,1 mm - 3 mm AW	1,1 mm - 2,0 mm AW	0 mm - 1 mm AW
3	Mehr als 1 cm AW	0,75 cm - 0,99 cm AW	0,51 cm - 0,75 cm AW	0,26 cm - 0,5 cm AW	Bis 0,25 cm AW
4	Mehr als 4 mm AW	3,1 mm - 4 mm AW	2,1 mm - 3 mm AW	1,1 mm - 2,0 mm AW	0 mm - 1 mm AW
5	Drei Phasen	-	Zwei Phasen	Eine Phase	Keine Phase
6	-	-	Dritter Platz	Zweiter Platz	Erster Platz

IV. EVALUATION

In diesem Kapitel findet die Bewertung anhand des Bewertungskatalogs durch das vorgestellte Rating Schema statt. Die gewonnen Werte entsprechen dem arithmetischen Wert der jeweiligen Messungen.

A. Bewertung durch Anzahl der Punkte

Da sich die Anzahl der Punkte der Punktwolke nur für einen relativen Vergleich eignen, wird das beste Ergebnis als Referenzwert verwendet.

TOF: Die Anzahl der durch den TOF-Ansatz erzeugten Punktwolken besteht aus 368.640 Punkten. Es gilt zu betonen, dass sämtliche Aufnahmen stets aus derselben Anzahl an Punkten besteht. Dieses Verhalten konnte für jedes Paket aus jeder Entfernung festgestellt werden. Dieses Ergebnis erwies sich als das beste unter den verglichenen Kamerasystemen, weshalb es als Referenzwert für der Bewertung dient. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.

Stereo: Die Anzahl der Punkte, aus denen sich die Punktwolke des Stereo-Ansatzes zusammensetzt, variiert in jeder Aufnahme. Die durchschnittliche Anzahl der enthaltenen Punkte beträgt 234.400. Die Punktwolke, welche die höchste Anzahl an Punkten aufweist, unterscheidet sich um 23341 Punkten zu der Wolke, welche die geringste Anzahl an Punkten aufweist. Hinsichtlich des Referenzwerts besitzt dieser Ansatz eine Reduktion von 36,41 %. Das Testergebnis wird im Vergleich zum Referenzwert mit einem befriedigenden Ergebnis in Höhe von drei Sternen bewertet.

Selfmade: Die Punktwolken, welche mithilfe der beiden Webcams erzeugt wurden, weisen die größten Schwankungen in der Anzahl der Punkte auf. Die durchschnittliche Punktwolke setzt sich bei diesem Vorgehen aus 149.185 Punkten zusammen. Jeder Aufnahme ist einer unterschiedlichen Anzahl an Punkten zu entnehmen. Hinsichtlich des Referenzwerts besitzt dieser Ansatz eine Reduktion von 59,53 %. Das Testergebnis wird im Vergleich zum Referenzwert mit einem mangelhaften Ergebnis in Höhe von zwei Sternen bewertet.

B. Bewertung der Wiederholgenauigkeit

Für die Bewertung der Wiederholgenauigkeit wurden wiederholt Bildpaare aufgenommen, die als Grundlage für einen direkten Vergleich dienen. Unter Verwendung des Softwaretools CloudCompare wurden die jeweiligen Punkte in Relation zum nächstliegenden Punk berechnet und ein Mittelwert aus jedem Punkt für die gesamten Aufnahmen gebildet.

TOF: Die durch den TOF-Ansatz generierten Punktwolken weisen keine Veränderung in der Eigenschaft wiederholte Bildaufnahme auf. durch Selbst das Übereinanderlegen der generierten Punktwolken visualisiert, dass keine Änderung oder Qualitätseinbußen wahrzunehmen sind. Die Abbildung 7 visualisiert den Zustand des Übereinanderlegens der Aufnahmepaare. Bei diesem Vorgehen wurde eine durchschnittliche Distanz der Punkte von 0,48 mm bei einer Standardabweichung von 0,48 mm festgestellt. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.



Abb. 7. TOF-Ansatz, Bewertungskriterium 2, Überlappen der Aufnahmen

Stereo: In der Abbildung 8 ist das Maß an Wiederholgenauigkeit der Stereokamera abgebildet. Die exemplarischen Aufnahmen von (a) bis (d) weisen visuell Unterschiede in der Dichten und der Qualitäten der Punktwolken auf. Bei der softwaregestützten Analyse der Punkte konnte diese Beobachtung allerdings nicht bestätigt werden. Bei diesem Vorgehen wurde eine durchschnittliche Distanz der Punkte von 0,68 mm bei einer Standardabweichung von 0,63 mm festgestellt. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.



Abb. 8. Stereo-Ansatz, Bewertungskriterium 2, Darstellung der Fehler

Selfmade: In Abbildung 9 wird veranschaulicht, dass sich die Punktwolken, welche mithilfe der beiden Webcams erzeugt werden, trotz derselben Szene und identischen Verhältnissen voneinander unterscheiden. Eine Reproduzierbarkeit ist demnach nicht gegeben. Bei diesem Vorgehen wurde eine durchschnittliche Distanz der Punkte von 5,21 mm bei einer Standardabweichung von 5,33 mm festgestellt. Das Testergebnis wird mit einem ungenügenden Ergebnis in Höhe von einem Stern bewertet.



Abb. 9. Selfmade-Ansatz, Bewertungskriterium 2, Darstellung der Fehler

C. Bewertung der relativen Messgenauigkeit

Ziel dieses Bewertungsaspektes ist es, die relative Messgenauigkeit der Punktwolken zu prüfen. Der Tabelle I sind die Ground-Truth Werte der einzelnen Pakete zu entnehmen. Es gilt allerdings zu betonen, dass die Messungen, die durch das Softwaretool Meshlab realisiert wurden, manueller Natur entsprechen. Daraus folgt, dass Abweichungen der Messwerte im Detail zu berücksichtigen sind. Ebenfalls bestehen Messungenauigkeiten beim Ermitteln der Ground-Truth da leichte Verformungen am Paket während des Versuchs aufgetreten sein könnten. Bei den in Tabelle III dargestellten Metriken wurde diese Ungenauigkeit berücksichtigt.

TOF: Die Abbildung 10 zeigt den Messprozess der Punktwolke, die dem TOF-Ansatz zugrunde liegen. In der Summe der Messungen wurden vereinzelt Unterschiede zur relativen Genauigkeit in Form von 0,1 cm festgestellt. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.



Abb. 10. TOF-Ansatz, Bewertungskriterium 3, Messung

Stereo: In der Abbildung 11 ist der Messprozess, dem der Stereo-Ansatzes unterliegt, dargestellt. In der Summe der Messungen wurden ein Unterschied in Form von 0,6 cm zur relativen Genauigkeit festgestellt. Das Testergebnis wird mit einem befriedigenden Ergebnis in Höhe von drei Sternen bewertet.



Abb. 11. Stereo-Ansatz, Bewertungskriterium 3, Messung

Selfmade: Der Messprozess der Punktwolke, welche mithilfe der beiden Webcams erzeugt wurde, ist in Abbildung 12 dargestellt. In der Summe der Messungen wurden Unterschiede zur relativen Genauigkeit in Form von 0,7 cm festgestellt. Das Testergebnis wird allerdings mit einem mangelhaften Ergebnis in Höhe von zwei Sternen bewertet, da die Aufnahmen Ausreiser enthielten, die eine Messung nicht möglich machten.



Abb. 12. Selfmade-Ansatz, Bewertungskriterium 3, Messung

D. Bewertung der Fähigkeit der Anpassung

Innerhalb des Bewertungskriteriums wurde der Einfluss von Störfaktoren wie Licht und das Vorhandensein von reflektierenden Oberflächen wie Klebeband auf die Qualität der Punktwolke getestet. Für die Bewertung wird dasselbe Schema wie bei der Wiederholgenauigkeit verwendet.

TOF: Die Abbildung 13 zeigt den Einfluss von erhöhten Lichtverhältnissen in Kombination mit dem Einsatz von Klebeband. Auf der rechten und linken Seite wurden Mobiltelefone installiert, die eine Veränderung der Lichtstärke erzeugen. Im direkten Vergleich zu den Paketen ohne Störfaktoren wurden keine Einschränkungen festgestellt. Die Berechnung der durchschnittlichen Distanz zwischen den Punkten liefert ein Ergebnis in Form einer Ungenauigkeit von 0,5 mm bei reflektierendem Material und 0,4 mm bei starken Lichtverhältnissen. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.



Abb. 13. TOF-Ansatz, Bewertungskriterium 4, Messung

Stereo: Der Abbildung 14 zeigt den Einfluss von Störfaktoren. Die Aufnahme (a) veranschaulicht den Einfluss von Klebeband auf das Paket, (b) den Einfluss des Störfaktors in Form von starkem Licht und (c) die Kombination aus beidem. Die Untersuchung hat ergeben, dass der Einfluss von reflektierendem Material zu einer Ungenauigkeit von 2,14 mm führt. Die Ergebnisse der Aufnahmen (b) und (c) weisen visuelle Löcher vor und auf den Paketen auf, die durch Einfluss der Lichtquelle entstanden sind. Allerdings wurde Messgenauigkeit nicht stark beeinflusst, da eine Ungenauigkeit von 0,76 mm festgestellt wurde. In Anbetracht der Ergebnisse der Wiederholgenauigkeit wurde eine geringfügige negative Auswirkung von 0,08 mm feststellen. Das Überprüfen mit beiden Störfaktoren liefert eine reflektionsdominierte Ungenauigkeit von 2,23 mm. Das Testergebnis wird mit einem befriedigenden Ergebnis in Höhe von drei Sternen bewertet.



Abb. 14. Stereo-Ansatz, Bewertungskriterium 4, Darstellung der Fehler

Selfmade: Die Abbildung 15 zeigt den Einfluss von Licht (b) und Klebeband (a) auf die Punktwolke, welche mithilfe von zwei Webcams erzeugt wurde. Die Berechnung der durchschnittlichen Distanz zwischen den Punkten liefert ein Ergebnis in Form einer Ungenauigkeit von 8 mm bei reflektierendem Material und bei starken Lichtverhältnissen. Das Testergebnis wird mit einem ungenügenden Ergebnis in Höhe von einem Stern bewertet.



Abb. 15. Stereo-Ansatz, Bewertungskriterium 4, Darstellung der Fehler

E. Bewertung der Kalt- und Warmlaufphasen

Innerhalb des Bewertungskriteriums soll erforscht werden, ob die Einsatzdauer einen Einfluss auf die Qualität der Punktwolken hat. Als Messgegenstand wurden hierfür die Kalt- und Warmlaufphase sowie der Normalzustand betrachtet.

TOF: Der Abbildung 16 ist in der Aufnahme (a) die Kaltlaufphase, in der Aufnahme (b) der aktive Normalzustand und in der Aufnahme (c) eine Warmlaufphase des TOF-Kamerasystems in Form von 8 Arbeitsstunden zu entnehmen. Im relativen Vergleich wurde in der Darstellung sowie der Dichte der Punktwolke keine Veränderung festgestellt. Die Qualität der Punktwolken bleibt konstant für den gemessenen Zeitraum von bis zu acht Stunden. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.





Stereo: Der Abbildung 17 ist derselbe Sachverhalt zu entnehme. Die Untersuchen haben ergeben, dass sich die Qualität der Punktwolke hinsichtlich des Normalzustands und der Warmlaufphase von bis zu acht Stunden konstant bleibt. In der Kaltlaufphase wurde festgestellt, dass die Qualität der Aufnahmen hinsichtlich der Anzahl der Punkte sowie der Genauigkeit nicht den gewohnten Ergebnissen entspricht. Nach einer Einsatzzeit von 15 Minuten nahm die Qualität wieder zu. Somit wurde eine Sensibilität in der Kaltlaufphase identifiziert. Das Testergebnis wird mit einem guten Ergebnis in Höhe von vier Sternen bewertet.



Abb. 17. Stereo-Ansatz, Bewertungskriterium 5, Darstellung der Fehler

Selfmade: Die Punktwolken, welche von den beiden Webcams erzeugt werden, variieren in ihrer Qualität von Beginn an sehr stark. Nach einer Nutzungszeit von 8 Stunden hat sich die Qualität der Punktwolken weder zum Besseren noch zum Schlechteren gewandelt. Somit weißt der Selfmade-Ansatz in jeder Phase unterschiedliche Ergebnisse auf. Das Testergebnis als nicht anwendbar betrachtet wird, erhält der Ansatz, den Durchschnitt der Gesamtpunktzahl in Form von zwei Sternen für diesen Bewertungspunkt, um das Ergebnis nicht zu verfälschen.

F. Bewertung der Kosten bzw. des Preises

Innerhalb dieses Kriteriums wurde ein Kostenvergleich der Kamerasysteme durchgeführt. Die Kosten der Kamerasysteme generieren sich aus den Anschaffungskosten sowie den laufenden Betriebskosten, welche aus dem Stromverbrauch der Kamerasysteme resultieren [14, pp. 366-367]. Angesichts des Anwendungsfalls der industriellen Versandstation wird davon ausgegangen, dass die Kamerasysteme 24 Stunden an 6 Tagen in 52 Wochen eingesetzt werden. Der Strompreis liegt bei Durchführung von diesem Kostenvergleich bei 43,9 Cent pro Kilowattstunde.

TOF: Der Listenpreis für das TOF-Kamerasystem von Microsoft liegt bei 410 \in . Der Stromverbrauch dieses Kamerasystem liegt bei 5,9 Watt, welches in unserem Anwendungsfall Stromkosten von 19,39 \in im Jahr verursacht. Im direkten Kostenvergleich ist die TOF-Kamera innerhalb der Anschaffungskosten und den Betriebskosten die teuerste Variante. Das Testergebnis wird mit einem befriedigenden Ergebnis in Höhe von drei Sternen bewertet.

Stereo: Der Listenpreis für das Kamerasystem von Intel liegt bei \$314 und umgerechnet 294,88 € entspricht. Der Stromverbrauch liegt bei 4,25 Watt die Stunde. In unserem Anwendungsfall werden somit Stromkosten in Höhe von 13,97 € verursacht. Im direkten Kostenvergleich schneidet die Intel Stereokamera in den Anschaffungskosten sowie in den Betriebskosten wesentlich günstiger ab als die TOF-Kamera. Das Testergebnis wird mit einem guten Ergebnis in Höhe von vier Sternen bewertet.

Selfmade: Der Listenpreis einer einzelnen Logitech c270 Webcam liegt bei jeweils 39,99 €. Die Anschaffungskosten belaufen sich somit auf 79,98 €. Der Stromverbrauch beider Webcams beträgt 4 Watt die Stunde. Es werden Stromkosten in Höhe von 13,14 € im Jahr für das Kamerasysteme anhand des Anwendungsfalls verursacht. Der Stromverbrauch zur Berechnung der Punktwolke floss nicht in die Stromkostenrechnung mit ein, da dieser nicht genau bestimmt werden konnte. Im direkten Kostenvergleich schneidet das Logitech c270 Stereokamerasystem als günstigste Variante ab. Die Anschaffungskosten von 79,98€ fallen dabei deutlich geringer aus als bei den anderen Kamerasystemen. Das Testergebnis wird mit einem sehr guten Ergebnis in Höhe von fünf Sternen bewertet.

G. Gesamtübersicht der Evaluierung

Die Tabelle IV zeigt die Bewertungen der einzelnen Kriterien der Kamerasysteme sowie das durchschnittliche auf eine Nachkommastelle gerundete Ergebnis.

Kriterium	TOF	Stereo	Selfmade
1	* * * * *	* * *	* *
2	* * * * *	* * * * *	*
3	* * * * *	* * *	* *
4	* * * * *	* * *	*
5	* * * * *	* * * *	* *
6	* * *	* * * *	* * * * *
Ø	4,7 *	3,7 *	2,2 *

 Tabelle IV.
 GESAMTÜBERSICHT DER EVALUIERUNG

V. FAZIT

Die Untersuchung hat gezeigt, dass sowohl das TOF-Vorgehen als auch das Stereo-Vorgehen gut für den untersuchten Anwendungsfall geeignet sind, da sie in der Lage sind präzise und zuverlässige Punktwolken zu generieren. Sofern der Kostenaspekt ein wesentlicher Indikator ist, kann auch der Einsatz die Tiefen-Stereokamera in Betracht gezogen werden. Im Vergleich dazu hat das Selfmade-Vorgehen jedoch nur eine geringe Anwendbarkeit gezeigt, da es zu ungenauen Ergebnissen führte. Hinsichtlich der eingangs vorgestellten Zielsetzung wurde festgestellt, dass sich das TOF-Vorgehen für den Anwendungsfall am besten eignet. Mit einem Gesamtergebnis von 4,7 wurden alle technischen Anforderungen zur vollsten Zufriedenheit umgesetzt. Die Tiefen-Stereokamera lieferte ein gutes Gesamtergebnis, das den wesentlichen Anforderungen entsprach. Allerdings ist in sämtlichen Bewertungskriterien ein Rauschen in den Punktwolken enthalten. Im Gegensatz zu den TOF- und dem Stereokamerasystem, welche sich durchaus in der Industrie für den Anwendungsfall eignen, haben die Untersuchungen ergeben, dass der Selfmade-Ansatz ungeeignet für den Anwendungsfall ist. Zudem gilt zu betonen, dass die Umsetzung durch diesen Ansatz in der Erstellung einer Punktwolke sowie in der Inbetriebnahme sehr komplex und störanfällig ist. Das Bewertungskriterium der Kosten sticht als einziges positiv hervor. Der Einsatz der Low-Budget-Methode kann für Lehr- oder Forschungsrahmen sinnvoll sein, um erste Erfahrungen in diesem Gebiet zu sammeln und neue Erkenntnisse zu treffen. Von dem Einsatz einer selbst gebauten Webcam basierten Anwendung im industriellen Bereich ist aufgrund der geringen Qualität abzuraten.

VI. KRITISCHE BETRACHTUNG

Es ist wichtig zu erwähnen, dass diese Studie als Einblick und richtungsweisend gewertet werden darf. In Anbetracht dessen, dass die Studie ausschließlich manuell evaluiert wurde, ist nicht auszuschließen, dass menschliche Fehler in die Bewertung eingeflossen sind. Besonders kritisch ist dies in Bezug auf die verwendeten Modellparameter, da diese nicht in einem statischen Labor, sondern in einem eigens erstellen Aufbau aufgestellt wurden. Leichte ungewollte Verrückungen können somit zu einem fehlerhaften Ergebnis führen. Die Abweichungen sind allerdings so gering, dass sie keine Auswirkungen auf das Ergebnis haben.

VII. AUSBLICK

Die Studie hat gezeigt, dass die gezielte Auswahl der domänenspezifischen Kamerasysteme für den Anwendungsfall einer Versandstation essenziell ist. Der Forschungstrend in diesem Bereich scheint in die Richtung von Kamerasysteme unter dem Einsatz von Maschine Learning zu gehen. Diese Studie liefert eine Grundlage für weitere Forschungsarbeiten im Sinne einer Weiterentwicklung mit dem TOF-Ansatz im Rahmen einer Roboteranwendung mit dem Fokus auf Kinematik.

QUELLEN

- R. Obermaier, Ed. Handbuch Industrie 4.0 und digitale Transformation: Betriebswirtschaftliche, technische und rechtliche Herausforderungen. Wiesbaden, Heidelberg: Springer Gabler, 2019. [Online]. Available: http://www.springer.com/
- [2] C. Poss et al., "Robust Framework for intelligent Gripping Point Detection," IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, 2019, pp. 717-723, doi: 10.1109/IECON.2019.8927308.
- [3] K. Mouri, H. Lu, J. K. Tan, and H. Kim, "Object Detection on Video Images Based on R-FCN and GrowCut Algorithm," in 2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), Busan, 2018, pp. 1–4, doi: 10.1109/ICT-ROBOT.2018.8549879.
- [4] R. Zeng, Y. Wen, W. Zhao, and Y.-J. Liu, "View planning in robot active vision: A survey of systems, algorithms, and applications," *Comp. Visual Media*, vol. 6, no. 3, pp. 225–245, 2020, doi: 10.1007/s41095-020-0179-3.
- [5] D. Le Hanh and C. V. Thien, "Visual guidance of a sealant dispensing robot for online detection of complex 3D-curve seams," *Int J Interact Des Manuf*, vol. 16, no. 4, pp. 1525–1532, 2022, doi: 10.1007/s12008-022-00843-y.
- [6] D. Le Hanh and K. T. G. Hieu, "3D matching by combining CAD model and computer vision for autonomous bin picking," *Int J Interact Des Manuf*, vol. 15, 2-3, pp. 239–247, 2021, doi: 10.1007/s12008-021-00762-4.
- [7] R. Szeliski, "Computer Vision: Algorithms and Applications, 2nd Edition,"
- [8] G. Anthonys, *Timing Jitter in Time-of-Flight Range Imaging Cameras*, 1st ed. (Springer eBook Collection). Cham: Springer International Publishing; Imprint Springer, 2022.
- [9] S. Giancola, M. Valenti, and R. Sala, A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies. Cham: Springer International Publishing, 2018.
- [10] L. Yang, I. Dryanovski, R. G. Valenti, G. Wolberg, and J. Xiao, "RGB-D camera calibration and trajectory estimation for indoor mapping," *Auton Robot*, vol. 44, no. 8, pp. 1485–1503, 2020, doi: 10.1007/s10514-020-09941-w
- [11] Fabian Sinz, "Kamerakalibrierung und Tiefenschätzung: Ein Vergleich von klassischer Bündelblockausgleichung und statistischen Lernalgorithmen,".
- [12] P. Suchde, T. Jacquemin, and O. Davydov, "Point Cloud Generation for Meshfree Methods: An Overview," *Arch Computat Methods Eng*, 2022, doi: 10.1007/s11831-022-09820-w
- [13] W.-C. Chang and C.-H. Wu, "Candidate-based matching of 3-D point clouds with axially switching pose estimation," *Vis Comput*, vol. 36, no. 3, pp. 593–607, 2020, doi: 10.1007/s00371-019-01642-5.
- [14] P. Römisch and M. Weiß, Projektierungspraxis Verarbeitungsanlagen: Planungsprozess mit Berechnung und Simulation der Systemzuverlässigkeit (Lehrbuch). Wiesbaden: Springer Vieweg, 2014.