

Multimodale Erfassung in Arbeitsräumen am Beispiel texturierter 3D-Innenrohroberflächen für die Kanalinspektion

Daniel Kapusi¹, Darko Vehar¹, Rainer Jahn¹, Rico Nestler^{1,2}, Karl-Heinz Franke¹

¹Zentrum für Bild- und Signalverarbeitung e. V.,
Werner-von-Siemens-Straße 10, D-98693 Ilmenau
eMail: {daniel.kapusi, darko.vehar, rainer.jahn, karl-heinz.franke}@zbs-ilmenau.de
URL: www.zbs-ilmenau.de

²TU Ilmenau, Fakultät für Maschinenbau, FG QBV,
Postfach 10 05 65, D-98684 Ilmenau
eMail: rico.nestler@tu-ilmenau.de, @zbs-ilmenau.de
URL: www.tu-ilmenau.de/qualitaetssicherung

Zusammenfassung. Der Beitrag widmet sich der Gewinnung multimodaler bildhafter Daten mit einem Kamera-Messkopf, der sowohl Farbbilder als auch 3D-Bilder nach dem Single-Shot-Prinzip mit hoher Genauigkeit und zeitlicher Verfügbarkeit erzeugen kann. In einer konkreten Umsetzung können mit einem Messkopfverbund, der auf einem speziell entwickelten Fahrwagen befestigt ist, während der Fahrt durch Abwasserkanäle kontinuierlich multimodale Bilder der Kanalinnenwand zur automatischen KI-basierten Analyse des Kanalschadenszustandes erfasst werden. Im Beitrag werden dazu grundsätzliche Überlegungen zur Konzeption des multimodalen Messkopfes und zur anwendungsspezifischen Umsetzung sowie die Schritte der sensornahen algorithmischen Verarbeitung der gewonnenen Bilddaten (Datenvorverarbeitung, 3D-Rekonstruktion, Registrierung von Textur- mit Tiefenbild sowie die 2D-3D-Registrierung zur Fusion von Bildern aus einem Messkopfverbund) vorgestellt. Im Beitrag wird weiterhin die Nutzbarkeit der Lösung in anderen Anwendungen, wie der Umfeldüberwachung auf mobilen Roboter-Plattformen in industriellen Arbeitsräumen, diskutiert.

1 Motivation

Im Zusammenhang mit aktuellen Anforderungen der Erfassung von 3D-Szenen in nichtstationärer Umgebung spielen u. a. die sogenannten Single-Shot-Verfahren eine zunehmende Rolle. Sie resultiert nicht zuletzt aus höheren lateralen Auflösungen bei Kameras und Projektoren, so dass heutzutage trotz lateraler Lichtkodierung gute 3D-Auflösungen erzielt werden können. Damit sind hohe Frameraten erzielbar, ohne Hochgeschwindigkeitskameras einsetzen oder übertriebenen Aufwand für die Dekodierung treiben zu müssen.

Für die dynamische Analyse von 3D-Szenen mit bewegten Objekten stehen beispielhaft Aufgaben der robusten Überwachung des Umfeldes auf mobilen Roboter-Plattformen für industrielle Arbeitsräume, der Mensch-Maschine- und der Roboter-Maschine-Kooperation oder der Qualitätssicherung mittels bewegter Sensorik. Zur

ersten Aufgabengruppe gehören Problemstellungen aus einem 2019 gestarteten Verbundprojekt RoboAssist (FKZ 03ZZ0448H), das im ZBS-Teilprojekt „Komponenten zur 3D-Umfelderfassung und –Interaktion in Arbeitsräumen“ u. a. mit ergänzender, kamerabasierter, intelligenter 3D-Erfassung des Umfelds einer autonomen mobilen Plattform befasst ist und sich in diesem Zusammenhang mit großen Arbeitsräumen, irritationsfreier latenzarmer Datenerfassung sowie echtzeitfähiger Objektverfolgung und –verortung auseinandersetzen wird. Zurückgegriffen werden soll dabei auf Erkenntnisse aus dem BMBF-Verbundprojekt AuZuKa („Automatisierte Zustandsanalyse von Abwasserkanälen durch virtuelle Begehung“ - FKZ 13N13895), das zur letzten der o. g. Aufgabengruppen gehört.

Im Mittelpunkt des Projektes AuZuKa stand die Entwicklung eines Inspektionsroboters, der während seiner Fahrt durch Abwasserrohre mit Durchmessern von 200 bis 400 mm 3D-Bilder der gesamten Kanalinnenwand aufzeichnet. Auf der Grundlage von fusionierten, farbtexturierten 3D-Daten der gesamten Kanalwand soll eine nachgelagerte, automatische, KI-basierte Schadensdetektion Schäden, wie z. B. Risse und Rissbildungen, Wurzeleinwüchse, Muffenversätze, Abplatzungen, Ablagerungen, etc., erkennen, klassifizieren und verorten. Die Aufgaben des ZBS e. V. waren hierbei die Konzeption eines multimodalen 3D-Messkopfes sowie die Aufbereitung, Registrierung und Fusion der Rohbilddaten zu mit Tiefeninformation aufgewerteten Farbbildern (RGB-D-Bilder, 2D/3D-Stitching). Besondere Herausforderungen dieser Entwicklung waren neben der Gewährleistung von Bildern mit hoher zeitlicher Verfügbarkeit und hoher geometrischer Genauigkeit die Konzeption eines sehr kompakten Messkopfes für die kleinen Arbeits- und Aktionsräume des Inspektionsroboters.

Als Ergebnis dieser Entwicklung liegt nunmehr ein Funktionsmuster eines multimodalen 3D-Messkopfes vor, das sowohl aus Sicht der Kamera-Projektor-Module und des Modulverbundes im Messkopf als auch aus Sicht der Umsetzung der 3D-Erfassung nach dem Single-Shot-Prinzip mit lateralen Mustern neuartig ist. Insbesondere die vom ZBS e. V. spezifisch erarbeitete Multi-Messkopfanordnung zur vollständigen Rundum-Erfassung der Kanalinnenwand führt nach einer algorithmischen 3D-Fusion zu einer neuen Qualität von automatischen Bildauswerteprozessen. Aufgrund des fortgeschrittenen Standes im Projekt AuZuKa werden die Ergebnisse der Kanalinspektion in den Mittelpunkt dieses Beitrages gestellt, der sich wie folgt gliedert:

In Abschnitt 2 wird der 3D-Messkopf vorgestellt. Es folgt in Abschnitt 3 die Beschreibung der Prozesskette vom Eingang der Rohbilddaten bis zur Erzeugung eines farbtexturierten Tiefenbildes. Abschnitt 4 beschäftigt sich mit dem Stitching eines Messdaten. Hier wird zunächst die 2D/3D-Registrierung von benachbarten Messbereichen erläutert. Anschließend wird die Umsetzung der Datenfusion, die der Vervollständigung der Daten in lückenhaften Bereichen und der Minderung der Messunsicherheit dient, dargestellt. Der Beitrag schließt in Abschnitt 5 mit einer Zusammenfassung.

2 Multimodaler Messkopf

Die Grundelemente des Messkopfes für die Schadenserkenkung, Bewertung und Vermessung im Anwendungsfall Abwasserkanal sind Kamera-Projektor-Module gemäß Abb. 1. Ein solches Modul besteht im Wesentlichen aus einer Projektions- und einer Kameraeinheit zur räumlichen Erfassung von Oberflächen nach dem Prinzip

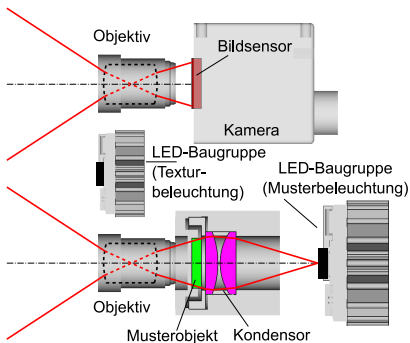


Abb. 1: Kamera-Projektor-Modul - Prinzipdarstellung (l), prototypischer Aufbau (r)

der aktiven Stereoskopie. Entsprechend einer Standard-Stereo-Anordnung enthalten Projektions- und Kameraeinheit gleiche Objektive und haben gleiche, parallel versetzte Abbildungsstrahlengänge. Das projizierte Muster besteht aus einem speziell lateral kodierten Binärmuster, dessen Kodewörter über ausreichend große Bereiche (Disparitätslimit) eindeutig sind. Zur Projektionsbeleuchtung dient eine Power-LED, deren Strahlung verlustarm in den Projektionsstrahlengang eingekoppelt wird. Zusätzliche LEDs an der Frontseite des Moduls realisieren eine diffuse Lichtquelle zur Aufnahme gewöhnlicher Farbbilder der Oberfläche (Texturbilder). Das Modul wurde für Nahaufnahmen und große Feldwinkel konzipiert, was an die Optiken besondere Anforderungen stellt. Der Inspektionsroboter nimmt die Daten zyklisch während seiner kontinuierlichen Fahrt durch das Kanalnetz auf. Die Ansteuerung der Sensorik und der Beleuchtung sind so ausgelegt, dass eine großzügige Überlappung der Bilder in Fahrtrichtung gegeben ist, um ein robustes Stitching der Einzelaufnahmen in den darauffolgenden Verarbeitungsschritten zu gewährleisten.

Bereits ein einzelnes Modul ist in dieser oder anwendungsbezogen modifizierter Form für vielfältige Inspektionsaufgaben geeignet. Im o. g. Anwendungsfall war es jedoch erforderlich, mehrere zueinander gedrehte Module in einem Messkopf zu integrieren. Die Sichtfelder der Module sind von der Rohrmittelachse aus auf die Rohrinnenwand gerichtet und bilden in ihrer Gesamtheit den Rohrfumfang vollständig und überlappend ab, so dass angrenzende Bilder registriert und vereinigt werden können. Alle Kamera-Projektor-Module des Messkopfes wurden jeweils separat mit einem planaren Kalibriertarget und mit Hilfe des Toolkit 3D-EasyCalib™ [1] vom ZBS e. V. geometrisch kalibriert. Die gewonnenen Kalibrierparameter bleiben für unterschiedliche Rohrdurchmesser unverändert.

3 Erzeugung eines farbtexturierten Tiefenbildes (RGB-D)

Im Anwendungsfall Kanalrohr lässt sich das zur Erzeugung von Tiefendaten projizierte Muster spektral nicht vom eigentlichen Szeneninhalte separieren (Unmixing), wie das etwa über spektral spezifisch sensitive Kanäle möglich wäre. Somit kann in einer

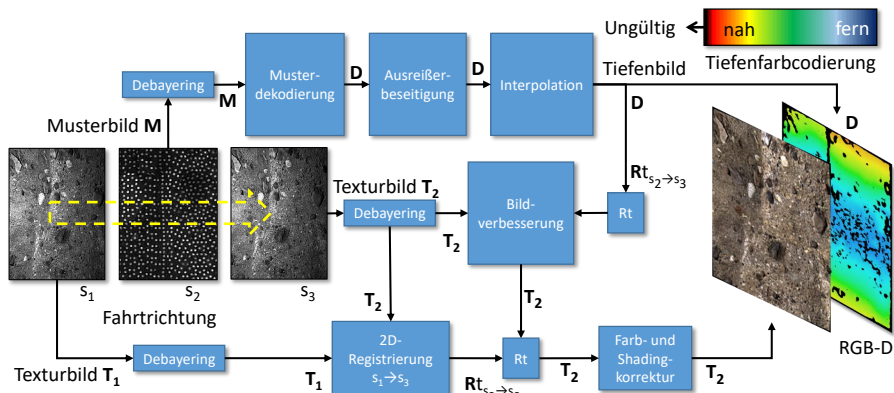


Abb. 2: Schrittfolge zur Erzeugung eines farbtexturierten Tiefenbildes (RGB-D)

simultan erzeugten Aufnahme das Muster nicht klar von der Farbtextur abgegrenzt werden, was die getrennte Farbtextur- und 3D-Auswertung stört. Aufgrund der Randbedingungen, dass Einflüsse des Umgebungslichts im Kanalrohr vernachlässigbar sind, die Kamerabewegung in den relevanten Zeitintervallen konstant und geradlinig erfolgt sowie der Szeneninhalt selbst quasi-statisch ist, kann die Aufnahme des Textur- und des Musterbildes mit ein und demselben Sensormodul zeitversetzt erfolgen.

In alternativen Szenarien mit willkürlich bewegten Messkopfsystemen, wie zum Beispiel auf einem mobilen Roboter und bei Szenen mit bewegten Objekten (RoboAssist), lässt sich ein zeitversetzter Ansatz zur Registrierung von Textur- und Musterbild nur bei vergleichsweise hohen Bildraten realisieren. Hinzu kommt, dass zum einen in der Regel Umgebungslicht die Musterprojektion überlagert und somit die Musterdetektion erschwert und zum anderen, dass die projizierten Muster störend auf Menschen im Umfeld wirken können und demzufolge irritationsfrei sein müssen. Diese Aspekte führen zu der Notwendigkeit, Muster- und Texturbild z. B. durch spektrale Trennung (z. B. RGB vs. NIR) tatsächlich simultan aufzunehmen. Praktisch kann dies über zwei getrennte Sensoren (z. B. RGB-Kamera und NIR-Kamera) realisiert werden, die zu einander geometrisch kalibriert sein müssen. Mit Hilfe der Tiefeninformationen kann die Parallaxe korrigiert und Textur- und Tiefenbild vor der Verarbeitung miteinander registriert werden. Theoretisch ist eine simultane Bildaufnahme auch über eine gemeinsame optische Achse und eine Kanaltrennung direkt auf dem Sensorchip (RGB-NIR-Chip) oder auch mit Hilfe einer Mehrchip-Kamera möglich. Zum aktuellen Zeitpunkt sind jedoch keine geeigneten Lösungen bekannt.

In Abb. 2 sind die Verarbeitungsprozesse von den Rohdaten bis zum RGB-D-Bild für den Anwendungsfall Kanalrohr graphisch dargestellt. Da sich die Messkopf-Kameras auf dem Inspektionsroboter kontinuierlich bewegen, muss der dadurch resultierende Versatz korrigiert werden. Das Muster- und Texturbild unterscheiden sich in ihren Inhalten ganz wesentlich, so dass eine direkte Registrierung über beispielsweise korrelationsbasierte Ansätze nicht erfolgversprechend ist. Durch Aufnahme eines zusätzlichen Texturbildes ergibt sich eine Textur T_1 , Muster M , Textur T_2 -Bildsequenz. Der

örtliche Versatz zwischen den Aufnahmen liegt im Bereich von einigen Millimetern. Da die Geschwindigkeit der Vorwärtsbewegung des Inspektionsroboters innerhalb des eng begrenzten Zeitfensters der Bildaufnahmesequenz als näherungsweise konstant angenommen werden kann, wird die ermittelte euklidische Transformation zwischen den Texturbildern auf die zeitlich und somit auch örtlich genau in der Mitte liegende Position zwischen Musterbild \mathbf{M} und Texturbild \mathbf{T}_2 interpoliert. Aufgrund der sehr kurzen Weglänge der Kamera zwischen den Aufnahmen \mathbf{M} und \mathbf{T}_2 resultieren nur sehr geringe Parallaxen, so dass tiefenabhängige Änderungen im Versatz vernachlässigbar sind. Die Registrierung zur Ermittlung von Rotation und Translation zwischen \mathbf{T}_1 und \mathbf{T}_2 erfolgt in Anlehnung an den Ansatz von Tzimiropoulos et al. [2], welcher aufgrund von Korrelationen im Frequenzbereich sehr schnell und wegen seines global wirkenden Charakters sehr robust ist.

Parallel zur Registrierung der Muster- und Texturbilder werden aus den Musterbildern unter Einbeziehung der vorkalibrierten Parameter der Sensor-Beleuchtungs-Module nach Musterdekodierung und 3D-Rekonstruktion die Tiefendaten erzeugt. In dem entstandenen Tiefenbild \mathbf{D} werden anschließend Ausreißer entfernt und prinzipbedingte Lücken durch Interpolation [3] geschlossen. Bevor \mathbf{T}_2 mittels Rotation und Translation (Block „Rt“ in Abb. 2) auf \mathbf{D} abgebildet wird, erfolgt die Korrektur von Unschärfen, die aus der Bewegung während der Bildaufnahme entstehen. Unter Einbeziehung von Vorwissen über die Szene wird dafür eine tiefengestaffelte Entfaltung mittels parametrischem Wiener-Filter [4] genutzt. In den Texturbildern treten globale und lokale Farbverfälschungen sowie Shadingeffekte auf, die wegen der besonderen Abbildungsverhältnisse in der diskutierten Anwendung sehr gravierend sind. Sie müssen deshalb durch Anwendung eines speziell entwickelten Korrekturverfahrens [5] für \mathbf{T}_2 behoben werden. Nach allen genannten Verarbeitungsschritten bilden das Texturbild \mathbf{T}_2 und das Tiefenbild \mathbf{D} das resultierende multimodale RGB-D-Bild eines Messkopfes.

4 2D/3D-Stitching

Das Stitching von RGB-D-Daten erfordert zunächst eine Registrierung, nach der die Textur- und Tiefendaten benachbarter Messbereiche für die nachfolgende Messdatenfusion (siehe Abschnitt 4.2) optimal in einander überführt werden können.

Neben der Anwendung bei der Vermessung von Abwasserkanalrohren ist das Stitching auch in anderen Anwendungen erforderlich, wenn größere Raumbereiche durch sich überlappende Teilszenen erfasst werden müssen. Der für RoboAssist interessante Fall des Stitchings nicht statischer Szenen wird im Rahmen dieses Beitrages nicht weiter untersucht. Die Problematik ist nach wie vor Gegenstand der Forschung [6] [7] [8] [9]. In der Literatur werden zu diesem Thema Ansätze wie „3D motion field“ oder „scene flow estimation“ [10] referenziert.

4.1 Featurebasierte 2D/3D-Registrierung

In Abb. 3 ist der Verfahrensablauf der featurebasierten 2D/3D-Registrierung skizziert. Es liegen die RGB-D Daten zweier benachbarter Messregionen vor, welche sich um mindestens 50% überlappen. Zunächst müssen korrespondierende Feature-Punkte

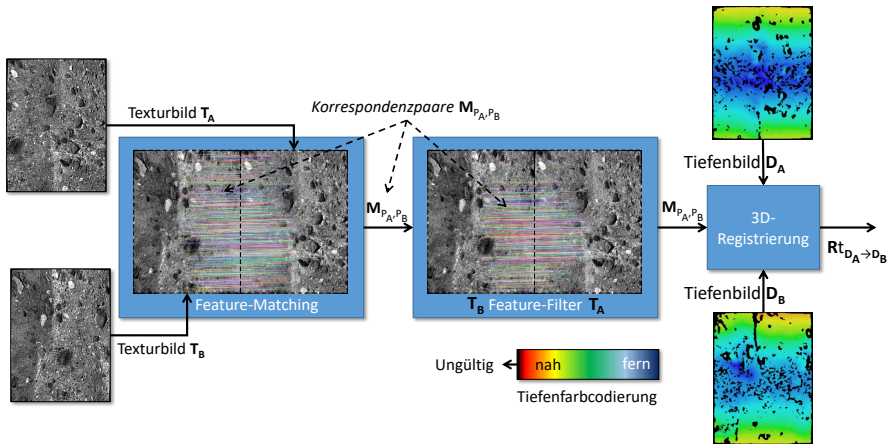


Abb. 3: Merkmalsbasierte, kamerabezogene 2D-/3D-Registrierung von Kanalinnenwandabschnitten

gefunden und miteinander gematched werden [11]. Sie werden in den Texturinformationen gesucht, die den Tiefendaten zugeordnet sind. Um den Suchbereich für das Matching möglichst weit einzuschränken, wird eine initiale Schätzung der Transformationsparameter durchgeführt, welche sich im Falle der Kanalinspektion auf die Lageinformationen eines Odometriemoduls stützt. Auf mobilen Roboterplattformen kann der enge Suchbereich für das Matching oftmals schon durch eine entsprechend hohe 3D-Datenrate erreicht werden, so dass die Datenpunkte von Bild zu Bild bereits nah beieinanderliegen. Die Liste mit kandidierenden Feature-Punkten wird anschließend nach einer ganzen Reihe von Kriterien schrittweise immer mehr ausgedünnt. Im letzten Filterschritt wird unter Einhaltung der Epipolarbedingung die Essentialmatrix geschätzt, aus der sich bereits die Rotation und der Richtungsvektor der Translation bestimmen lassen. Allerdings fehlt für die Translation noch der Betrag, d.h. die korrekte Skalierung. Für die Berechnung werden die 2D-Korrespondenzen aus den Texturbildern zusammen mit den Tiefenkarten und den bekannten intrinsischen Kameraparametern in 3D-Korrespondenzen überführt, so dass die euklidische Transformation analytisch eindeutig gelöst werden kann. Für die Bestimmung der Feature-Punkte sind eine Reihe von Verfahren an realen Messdaten getestet worden. Die klassischen SIFT-Merkmale [12] [13] lieferten bereits gute Ergebnisse. Aus Sicht der Rechte für eine praktische Umsetzung wurden lizenzrechtlich uneingeschränkt nutzbare Merkmale wie BRISK [14], ORB [15], KAZE [16] und A-KAZE [17] untersucht. In der vorliegenden Anwendung wurden Merkmale nach A-KAZE [17] aufgrund des Laufzeitverhaltens und der erreichbaren Güten für die Registrierung ausgewählt.

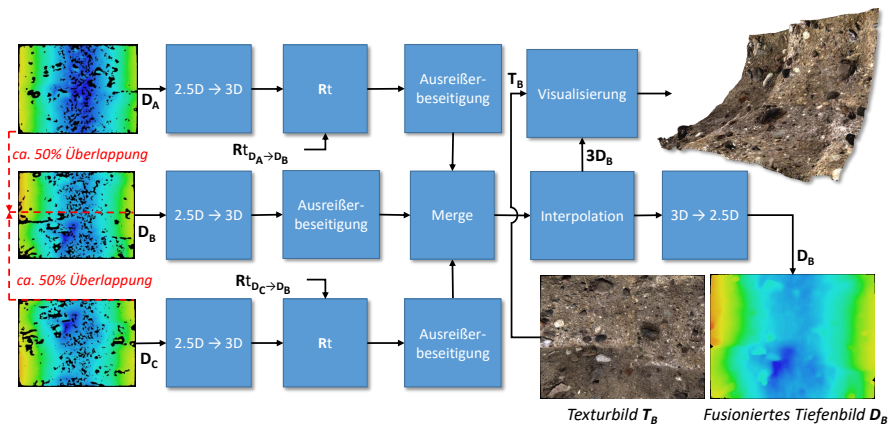


Abb. 4: 3D-Datenfusion

4.2 Datenfusion

Die Datenfusion dient dem Schließen von Lücken sowie der statistischen Minderung der Messunsicherheit durch Zusammenfassung korrespondierender Messpunkte benachbarter Messregionen in den Überlappungsbereichen.

Abb. 4 zeigt die schematische Darstellung der 3D-Datenfusion. Das Tiefenbild D_B enthält Lücken, die beispielsweise aufgrund von Abschattungen, extremen räumlichen Gradienten der Oberflächenregionen in Bezug auf Kamera und Projektor, ungünstigen Größenverhältnissen von Fehlerobjektausdehnung und Mustergröße sowie weiteren unkooperativen Abbildungssituationen zwischen projiziertem Muster, Objekt Oberfläche und Kamera entstehen können. Auch die unmittelbar benachbarten Tiefenbilder D_A und D_C enthalten Lücken, die sich jedoch aufgrund der jeweils verschiedenen Kamera-Projektor-Perspektiven in der Regel an unterschiedlichen Stellen befinden. Durch eine Verschmelzung der sich überlappenden Daten können die Lücken weitestgehend geschlossen und die Messunsicherheit kann in den bereits lückenlosen Bereichen aufgrund der Mehrfachmessungen gesenkt werden. Ausgangspunkt für die Datenfusion ist die erfolgreiche 2D/3D-Registrierung der bildhaften Messbereiche gemäß 4.1. Dann lassen sich sämtliche Messbereiche mittels euklidischer Transformation in ein gemeinsames Koordinatensystem übertragen. Nach einer Ausreißerkorrektur erfolgt der eigentliche Merge-Schritt. Zur Umsetzung wurde ein blockbasierter Ansatz entwickelt und implementiert. Alle Messpunkte des betrachteten Messbereiches B werden 2D-Rasterblöcken definierter Größe zugeordnet. Es erfolgt eine gutgewichtete Ausdünnung der Messpunkte eines jeden Blockes. Zur Sicherstellung minimaler Messpunktabstände wird der Vorgang für verschobene Rasterblöcke wiederholt. Anschließend werden die verbliebenen Punkte interpoliert und in ein reguläres Raster (Zieltiefenbild) abgebildet. Das Resultat ist ein fusioniertes RGB-D-Bild. Zu Visualisierungszwecken werden die Texturen auf die so fusionierten 3D-Daten gemappt (siehe Abb. 5, Abb. 6, Abb. 7).

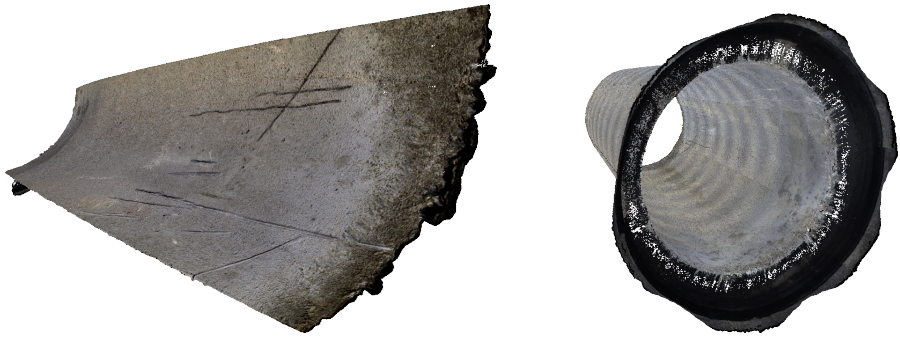


Abb. 5: Ausschnitt DN400 Betonrohr gestitched (l), komplett zusammengesetzt (r)



Abb. 6: Ausschnitt DN400 Betonrohr mit groben Gesteinseinschlüssen (gestitched)

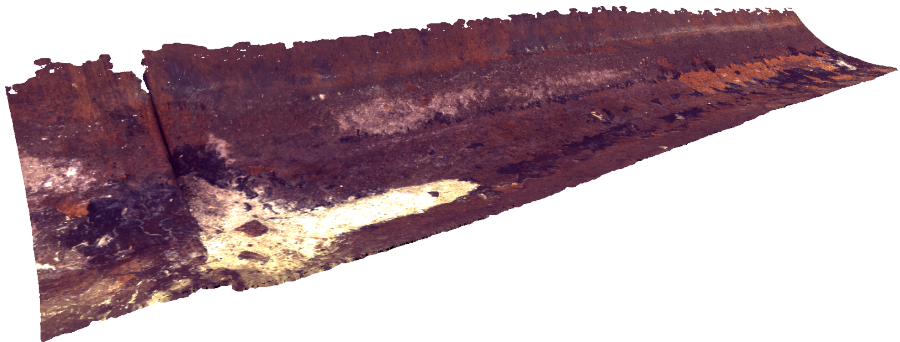


Abb. 7: Ausschnitt DN200 Steinzeugrohr mit Ablagerungen im Muffenbereich

5 Zusammenfassung

In dem hier vorliegenden Beitrag werden im Kern Ansätze zur multimodalen Erfassung von Kanalrohren mit Kamera-Projektor-Modulen nach dem Single-Shot-Prinzip vorgestellt. Dies umfasst die Beschreibung des optischen Aufbaus sowie der Verarbeitungsschritte zur Registrierung von Textur- und Tiefeninformationen und zum Stitching von Messdaten bei der 3D-Erfassung von Innenrohroberflächen. Die Kamera-Projektor-Messkopfmodule und die Methoden zur Verarbeitung sind vom Grundprinzip her auch auf andere Anwendungen, wie z. B. die Umfeldüberwachung auf mobilen Roboter-Plattformen in industriellen Arbeitsräumen, übertragbar. Allerdings stellen die wesentlich abweichenden Rahmenbedingungen (Raumbereiche, Umgebungslicht, Latenzzeiten, Frameraten, Analyseintervalle, Echtzeitforderungen) sowie gänzlich andere Zielstellungen neue Herausforderungen an hierfür einsetzbare ganzheitliche Lösungen. Erste Gedanken dazu wurden an geeigneten Stellen mit Bezug auf das Vorhaben RoboAssist ausgeführt.

Wir bedanken uns beim BMBF, das wesentliche Anteile der Arbeit im Rahmen der Verbundprojekte AuZuKa (FKZ 13N13895) sowie RoboAssist (FKZ 03ZZ0448H) fördert.

Literatur

- [1] D. Vehar, R. Nestler und K.-H. Franke, „3D-EasyCalib™ - Toolkit zur geometrischen Kalibrierung von Kameras und Robotern,“ in *22. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten, 3D-NordOst*, GFal Gesellschaft zur Förderung angewandter Informatik e. V., Dez. 2019.
- [2] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou und T. Stathaki, „Robust FFT-Based Scale-Invariant Image Registration with Image Gradients,“ *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jg. 32, Nr. 10, S. 1899–1906, 2010.
- [3] D. Shepard, „A Two-dimensional Interpolation Function for Irregularly-spaced Data,“ in *Proceedings of the 1968 23rd ACM National Conference*, Ser. ACM '68, New York, NY, USA: ACM, 1968, S. 517–524.
- [4] H. C. Andrews und B. R. Hunt, *Digital Image Restoration*. Prentice Hall Professional Technical Reference, 1977, ISBN: 0132142139.
- [5] D. Kapusi, R. Jahn, R. Nestler und K.-H. Franke, „Farb- und Shadingkorrektur im speziellen Einsatzfeld der automatischen Kanalinspektion,“ in *24. Workshop Farbbildverarbeitung*, Berlin: GFal Gesellschaft zur Förderung angewandter Informatik e. V., 2019.
- [6] T. Zhou, M. Brown, N. Snavely und D. G. Lowe, „Unsupervised Learning of Depth and Ego-Motion from Video,“ in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Juli 2017, S. 6612–6619.
- [7] X. Liu, M. Yan und J. Bohg, „MeteorNet: Deep Learning on Dynamic 3D Point Cloud Sequences,“ in *The IEEE International Conference on Computer Vision (ICCV)*, Okt. 2019.
- [8] W.-C. Ma, S. Wang, R. Hu, Y. Xiong und R. Urtasun, „Deep Rigid Instance Scene Flow,“ in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [9] A. Behl, D. Paschalidou, S. Donné und A. Geiger, „PointFlowNet: Learning Representations for Rigid Motion Estimation From Point Clouds,“ in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, S. 7962–7971.

- [10] Z. Lv, K. Kim, A. Troccoli, D. Sun, J. Rehg und J. Kautz, „Learning Rigidity in Dynamic Scenes with a Moving Camera for 3D Motion Field Estimation,“ in *European Conference on Computer Vision (ECCV)*, 2018.
- [11] M. Muja und D. G. Lowe, „Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration,“ in *VISAPP (1)*, INSTICC Press, 2009, S. 331–340, ISBN: 978-989-8111-69-2.
- [12] D. G. Lowe, „Distinctive Image Features from Scale-Invariant Keypoints,“ *International Journal of Computer Vision*, Jg. 60, Nr. 2, S. 91–110, Nov. 2004, issn: 0920-5691.
- [13] D. G. Lowe, „Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image,“ Pat. US6711293B1, 2000.
- [14] S. Leutenegger, M. Chli und R. Y. Siegwart, „BRISK: Binary Robust Invariant Scalable Keypoints,“ in *Proceedings of the 2011 International Conference on Computer Vision*, Ser. ICCV '11, Washington, DC, USA: IEEE Computer Society, 2011, S. 2548–2555, ISBN: 978-1-4577-1101-5.
- [15] E. Rublee, V. Rabaud, K. Konolige und G. Bradski, „ORB: An Efficient Alternative to SIFT or SURF,“ in *Proceedings of the 2011 International Conference on Computer Vision*, Ser. ICCV '11, Washington, DC, USA: IEEE Computer Society, 2011, S. 2564–2571, ISBN: 978-1-4577-1101-5.
- [16] P. F. Alcantarilla, A. Bartoli und A. J. Davison, „KAZE Features,“ in *Proceedings of the 12th European Conference on Computer Vision - Volume Part VI*, Ser. ECCV'12, Florence, Italy: Springer-Verlag, 2012, S. 214–227, ISBN: 978-3-642-33782-6.
- [17] P. F. Alcantarilla, J. Nuevo und A. Bartoli, „Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces,“ in *BMVC*, 2013.