
libBICOS: eine quelloffene Softwarebibliothek mit GPU-Beschleunigung für binäre Korrespondenzsuche zur 3D-Rekonstruktion

Christoph LIEBENDER, Michael BLEIER, Andreas NÜCHTER

*Photogrammetrie – Laserscanning – Optische 3D-Messtechnik
Beiträge der Oldenburger 3D-Tage 2025*

Zusammenfassung

Das folgende Forschungspapier präsentiert eine quelloffene Softwarebibliothek zur binären Korrespondenzsuche (BICOS), die effiziente 3D-Rekonstruktion mittels Stereokameras und strukturiertem Licht ermöglicht. Ausgehend von zwei Stapeln rektifizierter Stereobilder einer Szene die durch statistische Lichtmuster ausgeleuchtet wird, löst die Implementierung das Problem der pixelweisen Korrespondenzsuche. Unsere Implementierung mit GPU-Beschleunigung reduziert die Latenz der Tiefenbildberechnung auf weniger als 50 Millisekunden. Ausgehend vom durch DIETRICH ET AL. 2019 eingeführten BICOS-Deskriptor erweitern wir die Flexibilität des Algorithmus bzgl. anwendbarer Lichtprojektion und erhöhen die mögliche Deskriptorgröße. Schließlich geben wir mittels Laufzeitanalysen und Beispielrekonstruktionen einer Stereokamera einen Ausblick auf erreichbare Genauigkeiten und Performanz. Obwohl es viele Implementierungen zur Tiefenbildberechnung auf Basis eines einzelnen Bildpaares gibt, ist dies für Algorithmen auf Basis von Bildstapeln mit strukturiertem Licht nicht der Fall. Deshalb hoffen wir durch die Veröffentlichung unserer Bibliothek den Einstieg in das Thema zu erleichtern. Der Quellcode ist verfügbar unter <https://github.com/JMUWRobotics/libBICOS> unter der GNU LGPLv3.

1 Einleitung

Stereo Vision mithilfe von Kameras ist eine kosteneffiziente und vielseitige Lösung zur Rekonstruktion von 3D-Szenen. Üblicherweise wird das Prinzip der Stereorekonstruktion auf ein einzelnes, rektifiziertes Bildpaar zweier kalibrierter Kameras angewendet. Dabei wird durch Korrespondenzsuche in beiden Bildern ein Tiefenbild errechnet. Diese Variante kommt jedoch schnell an ihre Grenzen: unstrukturierte, unkooperative und nicht ausreichend beleuchtete Szenen erschweren das Finden von Korrespondenzen. Ein Ansatz zur Lösung dieses Problems ist die Projektion von zusätzlichen, strukturierten Lichtmustern, womit unstrukturierte Szenen rekonstruierbar gemacht werden.

Die Projektion von veränderlichen Lichtmustern ermöglicht Multi-Shot-Stereo Methoden, in denen der Intensitätsverlauf eines Grauwertpixels über mehrere Zeitpunkte hinweg im zweiten Bildstapel gesucht werden kann. Dieser Prozess der Korrespondenzsuche über zwei Bildstapel mit jeweils mehreren Bildern ist jedoch rechnerisch teuer: die Laufzeit von typischen

Korrelationsmethoden schränkt die Anwendbarkeit solcher Ansätze bei Echtzeitanwendungen ein.

Trotzdem bieten Methoden auf Basis der Korrelationsberechnung zweier Intensitätsverläufe genaue Ergebnisse. Idealerweise würde die Menge an Pixeln welche korreliert werden müssen auf eine Größenordnung kleiner als $O(HB^2)$ reduziert, mit Bildhöhe H und -breite B in Pixeln. Diese zu unterbietende Größenordnung ist in der epipolaren Geometrie von rektifizierten Stereosystemen begründet: für jeden Punkt in der linken Ansicht kommt eine Linie in der rechten Ansicht in Frage. Die implementierte binäre Korrespondenzsuche sucht dabei vorerst grob nach Korrespondenzen und reduziert damit die zu korrelierenden Intensitätsverläufe auf $O(HB)$.

Der Beitrag dieser Publikation ist:

- Eine ausführliche Beschreibung der Implementierungsdetails des BICOS-Algorithmus.
- Erweiterungen und Modifikationen des Originalansatzes (DIETRICH ET AL. 2019), um die Anwendbarkeit der Methode auf verschiedene Lichtmuster zu erhöhen.
- Eine Evaluation eines Bespielmessaufbaus bzgl. Genauigkeiten von 3D Rekonstruktionen.
- Die Veröffentlichung der vorgestellten Implementierung als frei verfügbare Softwarebibliothek.

2 Stand der Wissenschaft

Stereo Vision auf kurzen Baselines ist ein weit erforschtes Thema im Computer-Vision Bereich. Bezüglich Methoden in denen nur ein einzelnes Stereopaar verarbeitet wird ist Semi-global Matching (HIRSCHMÜLLER, 2008) eine weit verbreitete Vergleichsmethode. Ausgehend von pixelweiser Kostenaggregation über vertikale, schräge und horizontale Pfade im Bild und darauf folgende Disparitätsbestimmung durch Kostenminimierung berechnet Semi-global Matching dichte Tiefenkarten aus einzelnen Bildpaaren. Die Kostenberechnung kann dabei verschiedene Methoden einsetzen: Mutual Information (KIM ET AL., 2003), Census Transform (ZABIH UND WOODFILL, 1994) oder die Sample-Insensitive Metrik (BIRCHFIELD UND TOMASI, 1998).

Den State of the Art bzgl. einzelner Bildpaare nach den ETH3D (SCHOPS ET AL., 2017) bzw. KITTI (GEIGER ET AL., 2012) Challenges stellen erlernte Methoden dar: Neuronale Netzwerke die Features extrahieren um die Korrelation der Features entlang einer Bildpyramide zu bestimmen (LI ET AL., 2022), eine Kombination aus Einbild- und Zweibild Tiefenschätzung (LI ET AL., 2024) oder U-net Transformer (RONNEBERGER ET AL., 2015) die zusätzlich zur Tiefenkarte auch eine dazugehörige Konfidenzkarte ausgeben, um automatisierte Entscheidungsprozesse anhand der Tiefenkarte zu ermöglichen (MIN UND JEON, 2024). Trotz der Fortschritte im Bereich der KI-basierten Tiefenbildberechnung bleibt die Robustheit und Anwendbarkeit solcher Methoden auf realen Daten offen, da die eingesetzten Datensätze oft synthetisch generiert sind.

Anstelle der Optimierung von passiven Einzelbildpaarmethoden kann die Genauigkeit erhöht werden, indem ein in das Messsystem integrierter Projektor ein auf den Algorithmus abgestimmtes Lichtmuster projiziert. Dabei können Laserlinien (LIU ET AL., 2024; BLEIER UND NÜCHTER, 2017) zum Einsatz kommen, oder zweidimensionale Lichtmuster: Flecken- bzw. Punktmuster (GU ET AL., 2020; SCHAFFER ET AL., 2010; ZHONG ET AL., 2019), oder vertikales Streifenlicht (HEIST ET AL., 2016; SCHARSTEIN UND SZELISKI, 2003; ZHANG ET AL., 2018) welche einfach durch ein goes-before-optics (GOBO) Rad zwischen Projektor und Linsensystem erzeugt werden können. Veränderliche Lichtmuster legen Methoden nahe, Mehrbildpaarmethoden einzusetzen, in denen eine Zeitserie von linken und rechten Bildern zur Korrespondenzsuche genutzt wird (DIETRICH ET AL., 2019; HEIST ET AL., 2015). Schließlich können, wenn die Messlatenz keine Rolle spielt, vollständig kodierte Muster genutzt werden um jeden einzelnen Pixel einen eindeutigen Code zuzuweisen (BATLLE ET AL., 1998; SALVI ET AL., 2004; YOUNG ET AL., 2007).

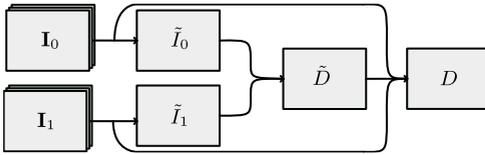


Abb. 1:

Flussdiagramm der Tiefenbildberechnung. Bildstapel I_i , Deskriptorbilder \tilde{I}_i , rohes Tiefenbild \tilde{D} , Tiefenbild D .

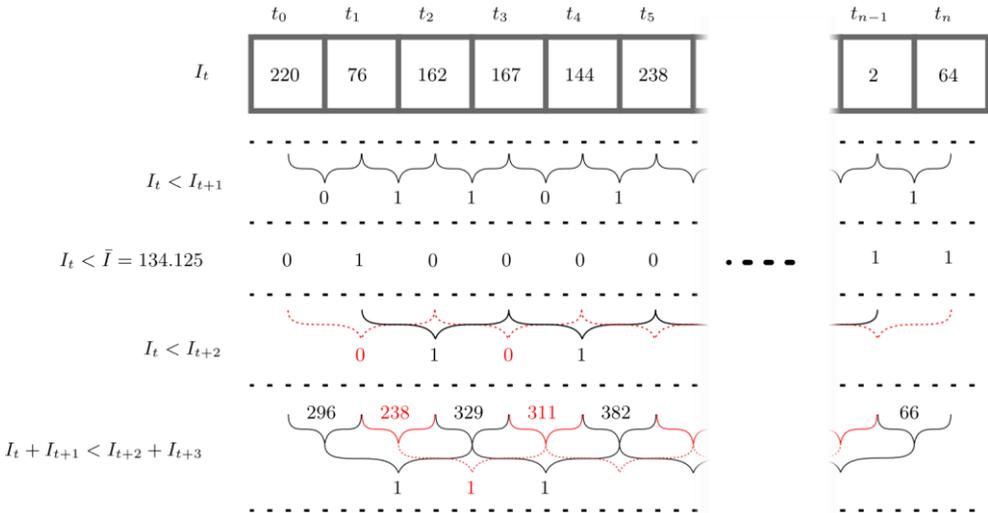


Abb. 2: Prinzip der Deskriptortransformation. Anhand einer Sequenz von Intensitäten I_t zu Zeitpunkten $t_{0...n}$ wird der Binärdeskriptor mittels Vergleichen zwischen diesen aufgebaut.

3 Algorithmus

Dieser Teil handelt von der Methodik hinter den in Abbildung 1 gezeigten Schritten. Zwei Bildstapel I_i werden in Deskriptorbilder \tilde{I}_i umgeformt. Diese werden dann erschöpfend anhand der epipolaren Geometrie miteinander verglichen um eine rohe Korrespondenz \tilde{d} im rohen Tiefenbild \tilde{D} zu bestimmen, wobei wir bereits einfache Konsistenzprüfungen verwenden. Schließlich prüfen wir Varianz und Korrelation der eigentlichen Intensitätsverläufe im Originalbildstapel anhand der rohen Korrespondenz um das resultierende Tiefenbild D zu erhalten.

3.1 Deskriptortransformation

Um die Eingabebildstapel zu Deskriptorbilder zu transformieren, wird jeder einzelne Intensitätsverlauf an den Pixeln einer Ansicht in einen Deskriptor transformiert. Hierfür orientieren wir uns an dem von DIETRICH ET AL. 2019 eingeführten, gleichnamigen Verfahren. Visualisiert in Abbildung 2 erhalten wir:

- $n - 1$ Bits von benachbarten Paaren I_t, I_{t+1} ;
- n Bits im Vergleich zum Intensitätsmittelwert \bar{I} ;
- $n - 2$ Bits von benachbarten Paaren mit einer Zwischenintensität I_t, I_{t+2} ;
- $n^2 - 5n + 6$ Bits von nichtüberlappenden Paarsummen.

Die letzten Bits werden durch den Vergleich zweier Summen gesetzt, deren Summanden aus benachbarten Intensitäten bestehen – mit der Einschränkung, dass die Summen aus verschiedenen Summanden bestehen. Die durch diesen letzten Teil generierten Bits sind quadratisch in n ; dadurch ist ihr Einfluss bei vielen Bildern vergleichsweise groß. Aus diesem Grund implementieren wir auch eine „limited“-variante für unsere Evaluation, in dem nur benachbarte, nicht-überschneidende Paarsummen verglichen werden. Dadurch reduzieren wir diesen letzten Teil auf $n - 4$ Bits.

DIETRICH ET AL. 2019 beschränken die Deskriptorgröße auf 64 Bits. Wir entfernen diese Beschränkung zunächst dadurch dass moderne Compiler mit `__uint128_t` einen Datentyp mit doppelter Anzahl an Bits unterstützen, und implementieren darüber hinaus Deskriptorgrößen von bis zu 256 Bits. Tatsächlich kürzen wir diese Deskriptorgröße zur Laufzeit auf 32- oder 64-Bit, wenn es die Größe des Eingabestapels erlaubt.

3.2 Epipolare Korrespondenzsuche

Ausgehend von beiden Deskriptorbildern suchen wir nun auf den epipolaren Linien nach pixelweisen Korrespondenzen. Für die Kostenfunktion richten wir uns nach DIETRICH ET AL. 2019 und berechnen die bitweise Hammingdistanz, definiert als die Anzahl der unterschiedlichen Bits zwischen beiden Deskriptoren:

$$C_{ham}(p, d) = \sum_j \tilde{I}_0(p, j) \oplus \tilde{I}_1(q, j) \quad (1)$$

Über j Bits in beiden Deskriptoren, wobei $q = (x - d, y)$. Moderne Prozessorarchitekturen reduzieren diesen Ausdruck auf zwei Instruktionen: `xor` und `popcount`. Die Disparität an einem Pixel p wird anschließend bestimmt durch:

$$d(p) = \operatorname{argmin}_d C_{ham}(p, d). \quad (2)$$

Eine große Anzahl an Pixeln pro Zeile resultiert in einer uneindeutigen Bestimmung dieser Disparität – schließlich können mehrere Deskriptoren den gleichen Hammingabstand haben. Um dem entgegenzuwirken, müssen die Korrespondenzen auf Konsistenz geprüft werden.

3.2.1 Globale Minima

Eine effiziente, weil einfach implementierte Methode zur Konsistenzprüfung bezeichnen wir hiernach als „Globale Minima“. Dabei wird erzwungen, dass ein Minimum nach (2) ein auf der epipolaren Linie globales sein muss – falls bei der epipolaren Suche mehrere Disparitäten den gleichen Wert nach (1) für einen Ausgangspunkt in der linken Ansicht erhalten, wird die Korrespondenz wegen Uneindeutigkeit verworfen.

3.2.2 Zurückmatchen

Eine aufwändigere, weil gründliche Variante stellt das „Zurückmatchen“ dar. Dabei wird für ein Punkt in der linken Ansicht ein Minimum in der rechten Zeile gesucht. Dieser Korrespondenzpunkt ist dann als Ausgangspunkt für eine Suche auf der linken Zeile zu verwenden. Ergibt die zweite Suche eine Korrespondenz die nicht gleich dem eigentlichen, linken Ausgangspunkt ist, wird die Korrespondenz verworfen. Um lokale Ähnlichkeiten zuzulassen verwenden wir in der Praxis ein Intervall um den linken Ausgangspunkt, das beim Zurückmatchen getroffen werden muss.

3.3 Korrelations- und Varianzfilter

Das so erhaltene Tiefenbild ist trotz erster Konsistenzprüfungen unbrauchbar, da die Hammingdistanz nach (1) nur die Anzahl der unterschiedlichen Änderungen kodiert und keine Information mehr über die zeitliche Anordnung dieser beinhaltet. Trotzdem haben wir nun eine initiale Tiefenschätzung, also eine eindeutige Zuordnung von linkem Punkt zu einem rechten Punkt. Um die resultierende Tiefenkarte zu erhalten können wir nun in die eigentlichen Intensitäten vergleichen. Hierfür nehmen wir die normalisierte Kreuzkorrelation und Varianzen als ihre Bestandteile:

$$NXC(p, d) = \frac{\operatorname{Cov}[I_0(p), I_1(q)]}{\sqrt{\operatorname{Var}[I_0(p)] \cdot \operatorname{Var}[I_1(q)]}}. \quad (3)$$

Dieser Korrelationswert kann nun mit einem Schwellwert $\theta_{NXC} \in (0,1)$ weiter gefiltert werden.

Zusätzlich zu diesem Kreuzkorrelationsfilter nutzen wir die dabei berechneten Varianzen um die Flexibilität von BICOS auf unterschiedlichen, gegebenenfalls suboptimalen Scanneraufbauten zu erhöhen. Dies wird insbesondere dann wichtig, wenn nicht der ganze Bildbereich von strukturiertem Licht abgedeckt ist – in so einem Fall entstehen durch Rauschen Korrespondenzen, die perfekt korreliert sind. Durch einen zweiten, beliebigen Schwellwert $\theta_{var} > 0$ entfernen wir Korrespondenzen, die nicht beleuchtet sind. Da die Berechnung von (3) die Berechnung der Varianzen bereits erfordert, entsteht dadurch kein Mehraufwand.

Im Gegensatz zu DIETRICH ET AL. 2019 setzen wir keinen Medianfilter ein, da unsere Nachverarbeitung in Kombination mit der höheren Deskriptorgröße bereits für ein rauscharmes Ergebnis ausreicht.

3.4 Subpixelinterpolation

Es ist klar, dass die durch ganzzahlige Pixelkoordinaten entstehende Diskretisierung die resultierende Tiefenkarte bzw. Punktwolke in Z-Richtung „abstuft“ und somit die Rekonstruktion von gekrümmten und geneigten Oberflächen verfälscht. Um diesem Effekt entgegenzuwirken, legen wir pro Bild ein Polynom zweiten Grades in die Umgebung einer konsistenten Korrespondenz, um das Maximum nach (3) auf diesen Polynomen als interpolierte Disparität zu bestimmen. (DIETRICH ET AL. 2019)

In der Praxis muss hierfür eine Schrittgröße gewählt werden; diese liegt in unseren Experimenten zwischen 0,01 und 0,5.



Abb. 3: Experimentelles Stereokamerasystem mit Projektor.

4 Experimente

Wir nutzen das Stereokamerasystem wie in Abbildung 3 dargestellt. Es besteht aus zwei synchronisierten 7MP ORX-10G-72S7C-C Kameras auf einem Abstand von 38cm. Anstelle eines Projektors mit GOBO-Rad zur Streifenlichtprojektion setzen wir einen handelsüblichen, unsynchronisierten Projektor ein. Dies belegt die Flexibilität unserer Implementierung. Die

Stereokameras werden mittels der Zhang'schen Methode (ZHANG, 2000) und 80 Bildpaaren eines ChArUco-Kalibrierungsmusters kalibriert. Wir behandeln in dieser Ausarbeitung ein Binärrauschen, welches an das statische Rauschen von Analogfernsehern erinnert. (Abbildung 4.) Andere Muster sind ebenfalls möglich (LIEBENDER ET AL., 2024), waren aufgrund der vergleichbar einfachen Erzeugung des Binärmusters jedoch nicht im Fokus unserer Experimente.

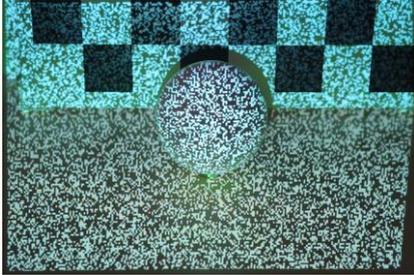


Abb. 4:

Projiziertes Binärrauschen.

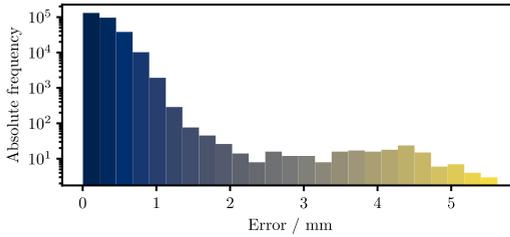
4.1 Freiformscans

Wir evaluieren drei Szenen: einen Gartenzwerger in 30cm Entfernung, für den ein Ground-Truth Scan vorliegt, eine Freiformszene „Adam & Eva“ auf 1m Entfernung und schließlich den Streifenlichtdatensatz von DIETRICH ET AL. 2019. Für die beiden letzteren liegt uns kein Vergleichsdatensatz vor. Trotzdem können wir qualitative Aussagen über die Ergebnisse unserer Implementierung treffen, auch weil die Qualität einer Rekonstruktion in großen Teilen von der Kalibrierung und nicht unbedingt nur vom eingesetzten Algorithmus abhängt. Für unsere Datensätze wenden wir die Konsistenzprüfung *Globale Minima* an, für den Streifenlichtdatensatz das *Zurückmatchen*.

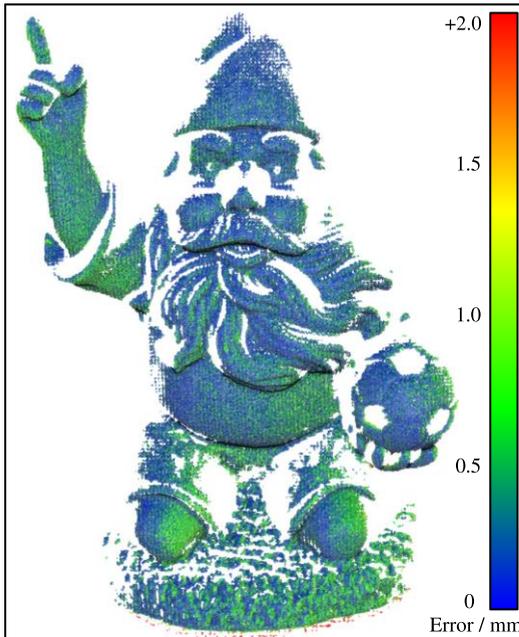


Abb. 5:

Der Gartenzwerger als Referenzobjekt. Links: Übersicht des Gartenzwergers, platziert auf einem Stuhl. Rechts: Ein- und Ausgabebildpaar. (Zur Darstellung wurde ein erhöhter Kontrast gewählt.)

**Abb. 6:**

Logarithmische Punkt- zu Punkt Fehlerverteilung aus Abbildung 7.

**Abb. 7:**

Auf das Ground-Truth Modell registrierter Gartenzwerg nach Segmentierung. Anhand der Cloud-to-Cloud Distanz eingefärbt mit Skala in Millimeter

4.1.1 Gartenzwerg

Abbildung 5 zeigt ein Eingabebildpaar, mit dem die Punktwolke aus Abbildung 7 berechnet wurde. Die Fehlerverteilung in Abbildung 6 zeigt, dass Punkt-zu-Punkt Fehler einen Millimeter nicht nennenswert übersteigen.

4.1.2 Adam & Eva

Eine größere Freiformszene ist die in Abbildung 8 dargestellte. Aufgrund ihrer Größe eignet sie sich gut, um die Anwendbarkeit unserer Implementierung auf größere Objekte zu überprüfen. Im Vergleich zu den Ergebnissen aus unserer vorherigen Publikation (LIEBENDER ET AL., 2024) haben Codeänderungen das Ergebnis der Subpixelinterpolation nochmals verbessert.

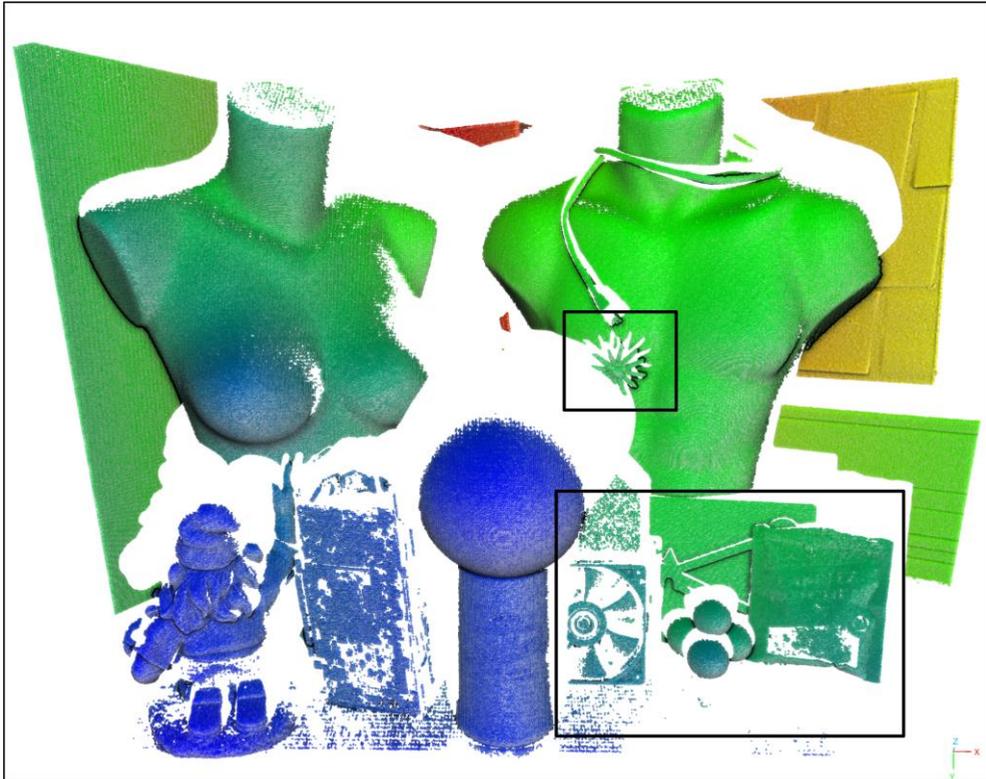


Abb. 8: Rekonstruierte Punktwolke der Freiformszene „Adam & Eva“, nach Tiefe eingefärbt. Oben: berechnete Punktwolke mit 33 Bildern und unserer De-

skriptorvariante. Unten links: Übersicht der Szene. Unten rechts: hervorgehobene Bereiche mit filigranen Details. Andere Parameter: $\theta_{NXC} = 0,96$, $\theta_{var} = 0,2$, $\sigma_s = 0,1$.

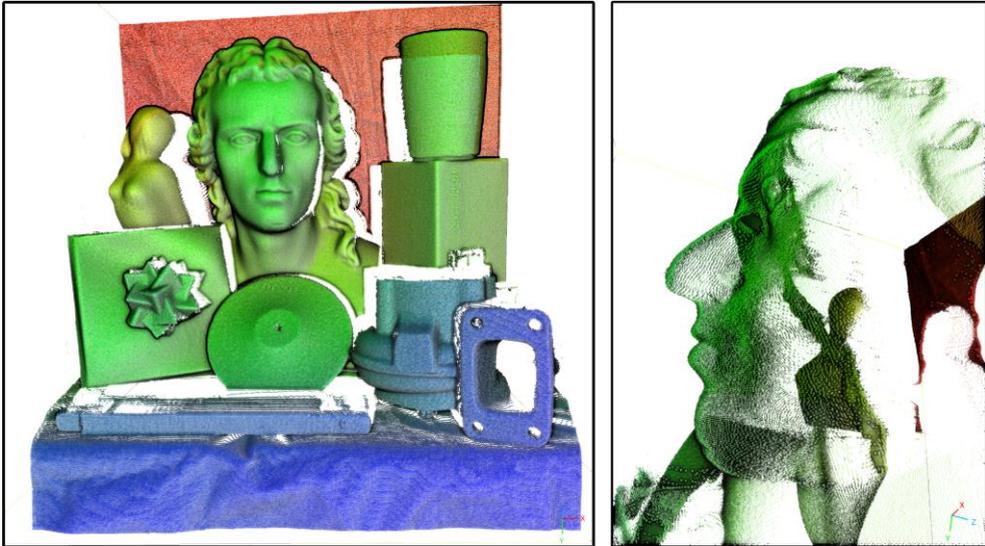


Abb. 9: Mit Konsistenzprüfung Zurückmatchen und sonst gleichen Einstellungen wie in Abbildung 8 berechnete Punktwolke aus dem Streifenlichtdatensatz von DIETRICH ET AL., 2019. Rechts ist zu erkennen, wie die Subpixelinterpolation Artefakte der „Tiefenstufen“ nahezu gänzlich entfernt.

4.1.3 Streifenlichtdatensatz

DIETRICH ET AL., 2019 stellen ihren Eingabedatensatz zur Verfügung, welcher mit einem Projektor mit GOBO Rad aufgenommene Bilder beinhaltet. Die unscharfen Lichtstreifen eignen sich gut, um den Effekt der Subpixelinterpolation darzustellen, da der tatsächliche Intensitätsverlauf gut durch ein quadratisches Polynom beschrieben werden kann. Da der Datensatz keine Vergleichscans beinhaltet, ist auch hier nur eine qualitative Aussage möglich. Abbildung 9 zeigt das von unserer Implementierung erzeugte Ergebnis.

4.2 Performance Benchmarks

Wir implementieren libBICOS in sowohl CUDA, also mit Grafikkartenbeschleunigung, als auch für Systeme ohne CUDA-Unterstützung in einem CPU-only Modus. Da letztere nicht optimiert ist beziehen wir uns für unsere Laufzeitmessung nur auf die CUDA-Implementierung. Dabei haben zwei Parameter einen maßgeblichen Einfluss auf die Laufzeit: die Anzahl der Bilder pro Eingabestapel und die Schrittgröße bei der Subpixelinterpolation. Die erzeugten Benchmarks zeigen libBICOS Version 1.2.2 auf einer Nvidia RTX 4090 und 3090.

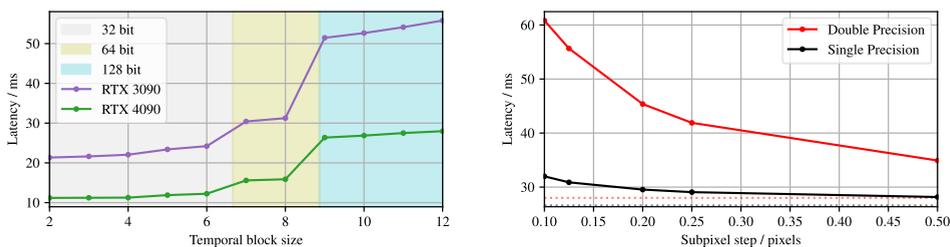


Abb. 10/11: Laufzeitanalyse von libBICOS bezüglich Bildstapelgröße (links) und Subpixelschrittgröße (rechts).

4.2.1 Bildanzahl

Abbildung 10 zeigt den naheliegenden Effekt der Bildanzahl auf die Laufzeit: mehr Bilder erhöhen diese. Da wir jedoch die Deskriptorgröße auf den nächstmöglichen Integerdatentyp reduzieren, entsteht ein Effekt der „Abstufung“: 2 bis 6 Bilder benötigen ~ 13 ms, 7 und 8 ~ 16 ms und darüber unter 30ms. Dieser Trend setzt sich bei Deskriptorgrößen größer als 128 Bit fort, ohne jedoch eine klare Abstufung zu zeigen; dies liegt daran, dass Deskriptoren größer als 64 Bit durch Bytearrays repräsentiert werden. Schließlich erhöht sich die Laufzeit auch dadurch, dass sich die Speicherzugriffszeiten erhöhen, weil für den Zugriff auf einen Intensitätsvektor alle Bilder in den Cache geladen werden müssen. Dieser Effekt wird besonders groß, wenn die Größe des Cache nicht ausreicht, um mehrere Bilder in diesem zu halten.

4.2.2 Subpixelschrittgröße

Die Subpixelschrittgröße gibt an, wie genau das quadratische Polynom um eine geschätzte Korrespondenz berechnet wird. Je größer diese wird, desto niedriger wird die Laufzeit, welche sich asymptotisch der Laufzeit ohne Subpixelinterpolation annähert. Abbildung 11 zeigt die Latenz der Disparitätsberechnung mit Subpixelinterpolation. Zusätzlich sehen wir dass die Reduktion von 64 Bit auf 32 Bit Genauigkeit bei der Fließkommaarithmetik die Berechnungszeit halbiert.

5 Fazit

In diesem Forschungspapier wurde die freie Implementierung eines Stereoalgorithmus basierend auf dem BICOS-Deskriptor vorgestellt. Durch GPU-Beschleunigung mittels CUDA erreichen wir eine Latenz zur Disparitätsberechnung von unter einer Zehntel Millisekunde. Wir erweitern den von DIETRICH ET AL. 2019 vorgestellten Algorithmus um variable Deskriptorgrößen von bis zu 256 Bit. Darüber hinaus stellen wir eine weitere, einfache Variante der Disparitätskonsistenzprüfung vor, die wir anstelle eines aufwändigeren Zurückmatchens einsetzen. Mit der zusätzlichen Voraussetzung eines Minimums an Varianz pro Intensitätsvektor erhöhen wir die Flexibilität der Methode, sodass 1. andere Lichtmuster angewendet werden können und 2. nicht das ganze Bild mit strukturiertem Licht ausgeleuchtet werden muss. Damit sinkt die Einstiegshürde – bisher bestehend aus komplexen Projektoraufbauten

– deutlich. Unsere Experimente zeigen, dass kalibrierte Stereokameras auf einer kurzen Baseline im Nahbereich Genauigkeiten von einem Millimeter erreichen und unterbieten. Indem wir unsere Implementierung unter der GNU Lesser General Public License (LGPLv3) frei zur Verfügung stellen, hoffen wir anderen Forschenden und der Industrie einen einfachen Einstieg in Mehrbildstereo mit strukturiertem Licht zu bieten.

Danksagung

Diese Arbeit wurde unterstützt durch das Projekt ‚Mixed Media Scanning‘ finanziert durch die Deutsche Forschungsgesellschaft (DFG NU230/14-1) und dem Projekt WISP finanziert durch das Bundesministerium für Wirtschaft und Klimaschutz (03EE3077D) aufgrund eines Beschlusses des deutschen Bundestages. Wir bedanken uns auch für die Unterstützung des Elitenetzwerks Bayern für den Studiengang Satellite Technology – Advanced Space Systems.

Literatur

- Dietrich, P., Heist, S., Landmann, M., Kühmstedt, P., & Notni, G. (2019). BICOS – An Algorithm for Fast Real-Time Correspondence Search for Statistical Pattern Projection-Based Active Stereo Sensors. *Applied Sciences*, 9(16), 3330.
- Hirschmuller, H. (2007). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2), 328-341.
- Kim, J. (2003, October). Visual correspondence using energy minimization and mutual information. In *Proceedings Ninth IEEE International Conference on Computer Vision* (pp. 1033-1040). IEEE.
- Zabih, R., & Woodfill, J. (1994). Non-parametric local transforms for computing visual correspondence. In *Computer Vision—ECCV'94: Third European Conference on Computer Vision Stockholm, Sweden, May 2–6 1994 Proceedings, Volume II 3* (pp. 151-158). Springer Berlin Heidelberg.
- Birchfield, S., & Tomasi, C. (1998). A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4), 401-406.
- Schops, T., Schonberger, J. L., Galliani, S., Sattler, T., Schindler, K., Pollefeys, M., & Geiger, A. (2017). A multi-view stereo benchmark with high-resolution images and multi-camera videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3260-3269).
- Geiger, A., Lenz, P., & Urtasun, R. (2012, June). Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 3354-3361). IEEE.

- Li, J., Wang, P., Xiong, P., Cai, T., Yan, Z., Yang, L., ... & Liu, S. (2022). Practical stereo matching via cascaded recurrent network with adaptive correlation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 16263-16272).
- Li, K., Wang, L., Zhang, Y., Xue, K., Zhou, S., & Guo, Y. (2024). LoS: Local Structure-Guided Stereo Matching. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 19746-19756).
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.
- Min, J., & Jeon, Y. (2024, October). Confidence Aware Stereo Matching for Realistic Cluttered Scenario. In 2024 IEEE International Conference on Image Processing (ICIP) (pp. 3491-3497). IEEE.
- Liu, Y., Ou, P., Xu, X., & Sun, J. (2024). Multi-line structured light binocular vision stereo matching method via coarse-to-fine spatial geometric constraints. *Optics & Laser Technology*, 176, 110950.
- Bleier, M., & Nüchter, A. (2017). Towards robust self-calibration for handheld 3D line laser scanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 31-36.
- Gu, F., Song, Z., & Zhao, Z. (2020). Single-shot structured light sensor for 3d dense and dynamic reconstruction. *Sensors*, 20(4), 1094.
- Schaffer, M., Grosse, M., & Kowarschik, R. (2010). High-speed pattern projection for three-dimensional shape measurement using laser speckles. *Applied optics*, 49(18), 3622-3629.
- Zhong, F., Kumar, R., & Quan, C. (2019). A cost-effective single-shot structured light system for 3D shape measurement. *IEEE Sensors Journal*, 19(17), 7335-7346.
- Heist, S., Lutzke, P., Schmidt, I., Dietrich, P., Kühmstedt, P., Tünnermann, A., & Notni, G. (2016). High-speed three-dimensional shape measurement using GOBO projection. *Optics and Lasers in Engineering*, 87, 90-96.
- Scharstein, D., & Szeliski, R. (2003, June). High-accuracy stereo depth maps using structured light. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. (Vol. 1, pp. I-I). IEEE.
- Zhang, C., Brahm, A., Breitbarth, A., Rosenberger, M., & Notni, G. (2018). Single-frame three-dimensional imaging using spectral-coded patterns and multispectral snapshot cameras. *Optical Engineering*, 57(12), 123105-123105.

- Heist, S., Lutzke, P., Dietrich, P., Kühmstedt, P., & Notni, G. (2015, June). Experimental comparison of laser speckle projection and array projection for high-speed 3D measurements. In *Optical Measurement Systems for Industrial Inspection IX* (Vol. 9525, pp. 282-289). SPIE.
- Battle, J., Mouaddib, E., & Salvi, J. (1998). Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern recognition*, 31(7), 963-982.
- Salvi, J., Pages, J., & Battle, J. (2004). Pattern codification strategies in structured light systems. *Pattern recognition*, 37(4), 827-849.
- Young, M., Beeson, E., Davis, J., Rusinkiewicz, S., & Ramamoorthi, R. (2007, June). Coded structured light. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). IEEE.
- Zhang, Z. (2002). A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11), 1330-1334.
- Liebender, C., Bleier, M., & Nüchter, A. (2024). libBICOS—An Open Source GPU-Accelerated Library implementing BInary COrrespondence Search for 3D Reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 57-64.

Liebender, Christoph

Julius-Maximilians-Universität Würzburg, Informatik XVII – Robotik

christoph.liebender@uni-wuerzburg.de

Bleier, Michael

Julius-Maximilians-Universität Würzburg, Informatik XVII – Robotik

michael.bleier@uni-wuerzburg.de

Nüchter, Andreas

Julius-Maximilians-Universität Würzburg, Informatik XVII – Robotik

Zentrum für Telematik e.V. – Würzburg

andreas.nuechter@uni-wuerzburg.de