

Verbesserung der Fähigkeiten von Vision-Sensoren

Ein Stitching-Algorithmus für 3D-Bilder für ein erweitertes Sichtfeld

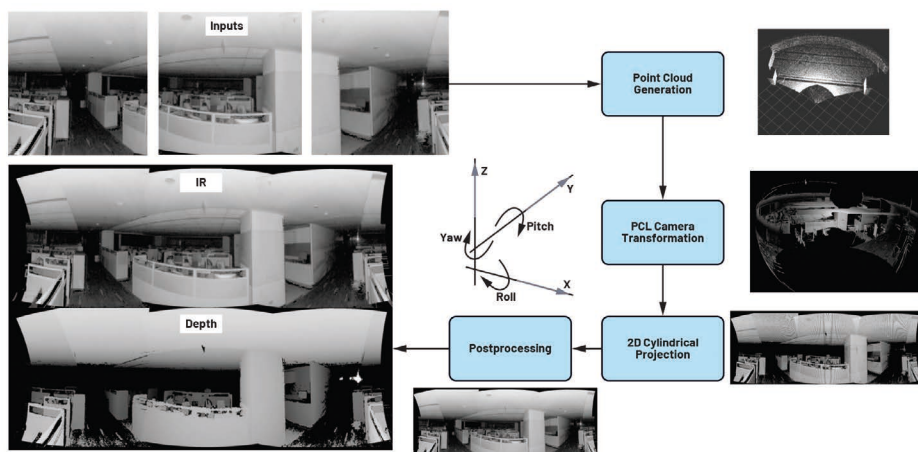


Bild 1: Ein Tiefen-Stitching-Algorithmus

Die zunehmende Beliebtheit von Time-of-Flight (ToF)-Kameras in industriellen Anwendungen, insbesondere in der Robotik, beruht auf ihren außergewöhnlichen Fähigkeiten in der Tiefenberechnung und Infrarot-Bildgebung. Trotz dieser Vorteile schränkt die mit dem optischen System verbundene Komplexität oft das Sichtfeld ein und begrenzt die Standalone-Funktionalität dieser Kameras. In diesem Artikel wird ein Stitching-Algorithmus zum Zusammenfügen von 3D-Bildern beschrieben, der für einen unterstützenden Host-Prozessor entwickelt wurde und die Notwendigkeit von Cloud-Berechnungen eliminiert. Dieser Algorithmus kombiniert nahtlos Infrarot- und Tiefendaten von mehreren ToF-Kameras in Echtzeit und erzeugt so ein kontinuierliches, hochwertiges 3D-Bild mit einem erweiterten Sichtfeld, das über das von Einzelgeräten hinausgeht. Die zusammengeführten 3D-Daten ermöglichen die Anwendung modernster Deep-Learning-Netzwerke, was insbesondere bei mobilen Roboteranwendungen von großem Nutzen ist, um die Visualisierung und Interaktion mit der 3D-Umgebung zu revolutionieren.

Definition Time-of-Flight-Kameras

Time-of-Flight (ToF)-Kameras sind herausragende Bildgebungssysteme mit außergewöhnlicher Reichweite, die ToF-Techniken verwenden, um die Entfernung zwischen einer Kamera und jedem einzelnen Punkt in einem Bild zu bestimmen. Dies erfolgt durch die Messung der Laufzeit eines künstlichen Lichtsignals, das von einem Laser oder einer LED ausgesendet wird. ToF-Kameras liefern präzise Tiefeninformationen und sind daher wertvolle Werkzeuge für Anwendungen, bei denen genaue Entfernungsmessungen und 3D-Visualisierungen entscheidend sind, wie z. B. in der Robotik und in der Industrietechnik, einschließlich der Kollisionserkennung und der Personenerkennung über ein Sichtfeld (FOV) von 270° zu Sicherheitszwecken.

Mit dem ToF-Sensor ADTF3175 lässt sich ein kalibriertes Sichtfeld von 75° erreichen. Allerdings ergeben sich Probleme, wenn das Sichtfeld einer Anwendung über diese Region hinausgeht und mehrere Sensoren erforderlich sind. Die Integration von Daten einzelner Sensoren zur Bereitstellung umfassender Analysen für die Gesamtansicht kann Schwierigkeiten bereiten.

Eine mögliche Lösung besteht darin, Sensoren einen Algorithmus auf einem Teilsichtfeld ausführen zu lassen und das Ergebnis zum Abgleich an einen Host zu übertragen. Dieser Ansatz ist jedoch mit Problemen wie Überlappungszonen, Funklöchern und Kommunikationslatenzen verbunden, was eine effektive Lösung zu einem komplexen Problem macht.

Stitching-Algorithmen

Ein alternativer Ansatz besteht darin, die erfassten Daten aller Sensoren zu einem einzigen Bild zusammenzufügen und anschließend Erkennungsalgorithmen auf das zusammengeführte Bild anzuwenden. Dieser Prozess kann auf einen separaten Host-Prozessor ausgelagert werden, wodurch die Sensoreinheiten von der Rechenlast befreit werden und Platz für erweiterte Analysen und andere Verarbeitungsoptionen geschaffen wird. Es ist jedoch wichtig zu beachten, dass herkömmliche Stitching-Algorithmen für Bilder von Natur aus komplex sind und einen erheblichen Teil der Rechenleistung des Host-Prozessors beanspruchen können. Darüber hinaus ist das Senden und Einfügen in die Cloud in vielen Anwendungen aus Datenschutzgründen nicht möglich.

Tiefen-Stitching-Algorithmus

Mit einer neuen algorithmischen Lösung können die IR- und Tiefenbilder der verschiedenen Sensoren unter Verwendung der Punktwolkenprojektionen der Tiefendaten zusammengefügt werden. Dazu werden die erfassten Daten mithilfe der extrinsischen Kamerapositionen transformiert und in den 2D-Raum zurückprojiziert, was ein einzelnes kontinuierliches Bild ergibt.

Dieser Ansatz erfordert nur minimalen Rechenaufwand, wodurch Echtzeit-Betriebsgeschwindigkeiten im Edge-Bereich erzielt werden und sichergestellt wird, dass die Rechenkapazität des Host-Prozessors für andere erweiterte Analysen verfügbar bleibt.

Autoren:
Rajesh Mahapatra
Senior Manager,
Anil Sripadarao
Principal Engineer,
Swastik Mahapatra
Senior Engineer
Analog Devices
www.analog.com

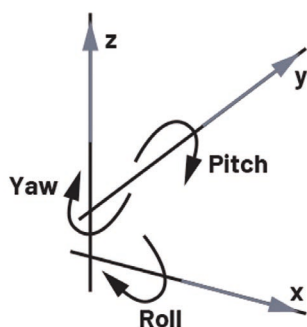


Bild 2: Extrinsische Kameraparameter



Bild 3: Eine zusammengeführte Punktwolke

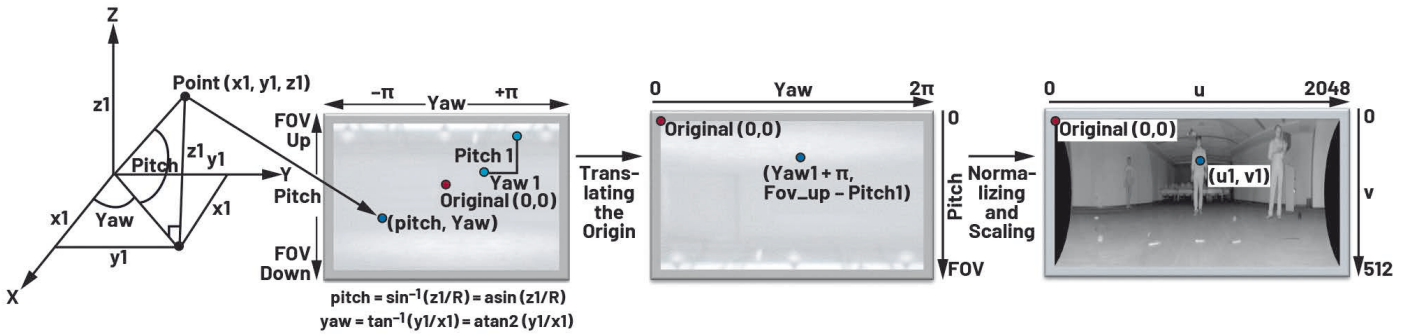


Bild 4: Ein Algorithmus für die zylindrische Projektion

Beschreibung der Lösung

Die 3D-ToF-Lösung von ADI besteht aus vier Phasen (Bild 1):

- **Vorverarbeitung der IR- und Tiefendaten:** Zeitsynchronisierung und Vorverarbeitung der IR- und Tiefendaten.
- **Projektion der Tiefendaten in eine 3D-Punkt- wolke:** Verwendung der intrinsischen Kamera- parameter zur Projektion der Tiefendaten in eine 3D-Punkt- wolke.
- **Transformation und Zusammenführung der Punkte:** Transformation der Punkte unter Verwendung der extrinsischen Kameraposi- tionen und Zusammenführung überlappender Bereiche.
- **Projektion der Punktwolke auf ein 2D-Bild:** Verwendung der zylindrischen Projektion, um die Punktwolke wieder auf ein 2D-Bild zu projizieren.

Im Folgenden sind die Herausforderungen und Lösungen für Systeme und Algorithmen beschrieben.

Empfang von IR- und Tiefenbildern durch den Host

Mehrere ToF-Sensoren sind über eine Hoch- geschwindigkeitsverbindung wie beispielsweise USB mit einem Host-Computer verbunden. Der Host erfasst Tiefen- und IR-Bilder und speichert diese in einer Warteschlange.

Synchronisation der IR- und Tiefendaten

Die vom Host empfangenen Tiefen- und Infra- rotbilder der einzelnen Sensoren werden zu unter- verschiedenen Zeitpunkten erfasst. Um zeitliche Unstimmigkeiten aufgrund der Bewegung von Objekten zu vermeiden, müssen die Eingaben aller Sensoren auf dieselbe Zeit synchronisiert werden. Dazu wird ein Zeitsynchronisationsmodul verwendet, das eingehende Frames anhand der Zeitstempel aus der Warteschlange abgleicht.

Projektion auf Punktwolken

Die Punktwolke wird auf dem Host unter Ver- wendung der synchronisierten Tiefendaten für jeden Sensor generiert. Jede Punktwolke wird dann basierend auf ihrer jeweiligen Kameraposi- tion (Bild 2) in der realen Welt transformiert (verschoben und gedreht). Anschließend werden diese transformierten Punktwolken zu einer einzigen kontinuierlichen Punktwolke zusam- mengeführt, die das kombinierte Sichtfeld der Sensoren abdeckt (Bild 3).

Projektion von 3D auf 2D

Die zusammengeführte Punktwolke des Sicht- felds wird mithilfe eines zylindrischen Projekti- onsalgorithmus, der auch als Frontsicht-Auf- projektion bekannt ist, auf eine 2D-Ebene pro- jiziert (Bild 4). Mit anderen Worten projiziert der Algorithmus jeden einzelnen Punkt der zusam- mengeführten Punktwolke auf ein Pixel in der 2D-Ebene, was zu einem einzigen kontinuier- lichen Panoramabild führt, das das kombinierte

Sichtfeld aller Sensoren abdeckt. Dies ergibt zwei zusammengefügte 2D-Bilder: eines für die zusammengefügte IR-Bilder und eines für die zusammengefügte Tiefenbilder, die jeweils auf eine 2D-Ebene projiziert werden.

Verbesserung der Projektionsqualität

Die Projektion der kombinierten 3D-Punkt- wolke auf ein 2D-Bild ergibt allerdings immer noch keine qualitativ hochwertigen Bilder. Die Bilder weisen Verzerrungen und Bildrauschen auf. Dies beeinträchtigt die optische Qualität und würde sich zudem negativ auf jeden Algorithmus auswirken, der auf die Projektion angewendet wird. Die drei Hauptprobleme (Bild 5) und deren Lösungen sind in den nachfolgenden Abschnit- ten dokumentiert.

Projektion ungültiger Tiefenbereiche

Die Tiefendaten des ADTF3175 liefern einen ungültigen Tiefenwert von 0 mm für Punkte, die außerhalb des Betriebsbereichs des Sensors (8000 mm) liegen. Dies führt zu großen leeren Bereichen im Tiefenbild und zu unvollständigen Punktwolken. Allen ungültigen Punkten auf dem Tiefenbild wurde ein Tiefenwert von 8000 mm (die größte von der Kamera unterstützte Tiefe) zugewiesen und damit eine Punktwolke erzeugt. Dadurch wurde sichergestellt, dass die Punkt- wolke keine Lücken aufweist.

Ausfüllen nicht zugeordneter Pixel

Bei der Projektion der 3D-Punktwolke auf eine 2D-Ebene entstehen nicht zugeordnete bzw. nicht

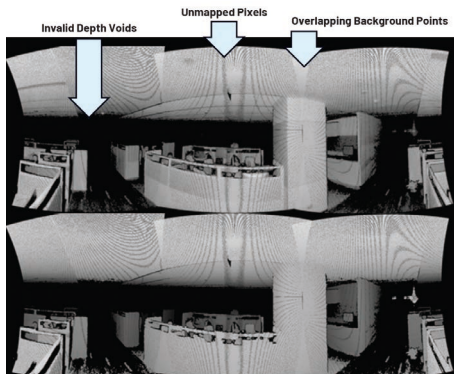


Bild 5: Probleme bei der 2D-Projektion

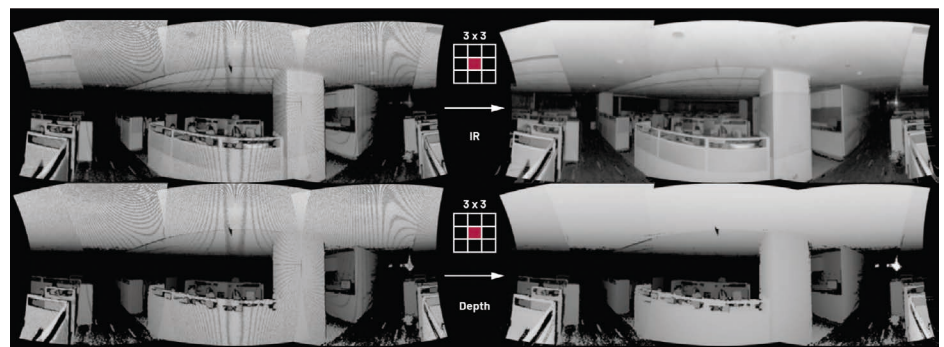


Bild 6: Ausfüllen nicht zugeordneter Pixel

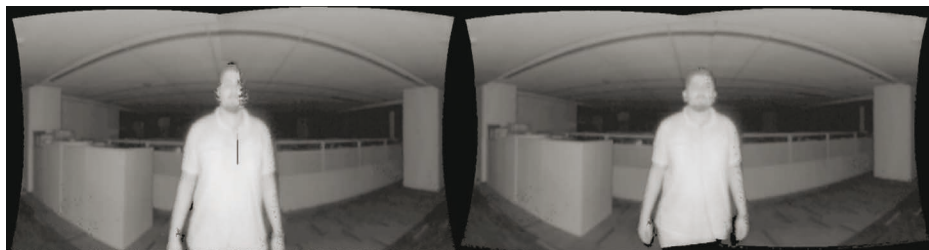


Bild 7: Behebung des Überlappungsrauschens



Bild 8: Zusammengesetzte IR-Daten mit einem Sichtfeld von 210°

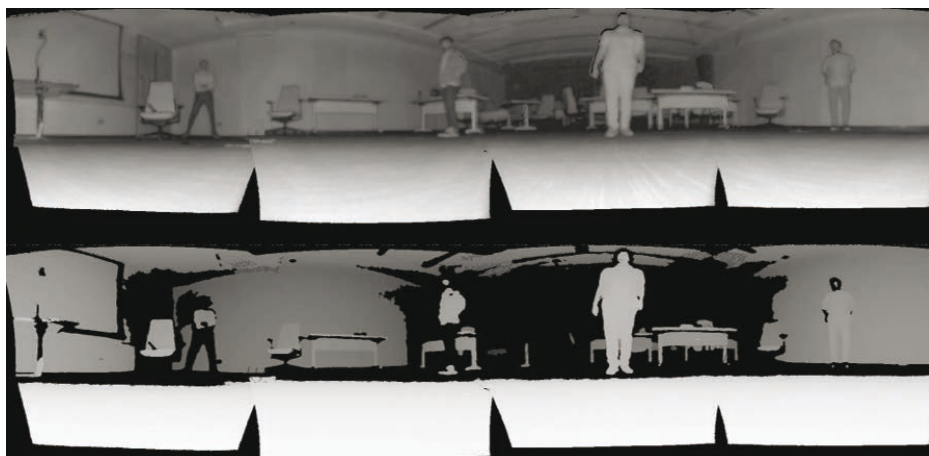


Bild 9: Zusammengesetztes IR- und Tiefenbild mit 278° Sichtfeld

ausgefüllte Bereiche im 2D-Bild. Viele Punktewolkenpixel (3D) werden demselben 2D-Pixel zugeordnet, sodass mehrere 2D-Pixel leer bleiben. Dies führt zu dem in Bild 6 gezeigten Dehnungsmuster. Um dieses Problem zu beheben, wurde ein 3×3 -Filter verwendet, der die nicht zugeordneten Pixel mit dem durchschnittlichen IR-/Tiefenwert der benachbarten 8 Pixel mit gültigen Werten füllt. Dies führte zu einem vollständigeren Ausgabebild sowie zur Beseitigung von Artefakten (Bild 6).

Bildrauschen durch überlappende Punkte

Der zylindrische Projektionsalgorithmus führt dazu, dass viele Punkte im überlappenden Bereich in der 2D-Projektion dieselben Zielkoordinaten erhalten. Dies erzeugt Bildrauschen, da die Hintergrundpixel die Vordergrundpixel überlappen. Um dies zu beheben, wird der radiale Abstand eines jeden Punktes mit dem vorhandenen Punkt verglichen, wobei der Punkt nur dann ersetzt wird, wenn sein Abstand zum

Kameraursprung kleiner ist als der des vorhandenen Punktes. Dadurch werden nur die Vordergrundpunkte beibehalten und die Projektionsqualität wird verbessert (Bild 7).

Schlussfolgerung

Dieser Algorithmus kann Bilder von verschiedenen Kameras mit weniger als 5° Überlappung zusammenfügen, im Gegensatz zu herkömmlichen Algorithmen, die eine Überlappung von mindestens 20° benötigen, um Schlüsselpunkte abzugleichen. Dieser Ansatz erfordert nur sehr wenig Rechenaufwand, was ihn zu einem idealen Kandidaten für Edge-Systeme macht.

Algorithmus	Durchschnittliche Gleitkommaoperationen
Herkömmliches Image-Stitching	857 Millionen
Vorgestelltes PCL-Tiefen-Stitching	260 Millionen (3,29-fache Verringerung)

Tabelle 1: Vergleich der Rechenkomplexität: Herkömmliche Algorithmen vs. vorgestellter Algorithmus für ein 512×512 QMP-Eingangsbild

Die Integrität der Tiefendaten bleibt nach dem Zusammenfügen erhalten, da es nicht zu Bildverzerrungen kommt. Diese Lösung unterstützt darüber hinaus die modulare Implementierung der ADTF3175-Sensoren, um das gewünschte Sichtfeld mit minimalem Verlust zu erhalten.

Die Erweiterung des Sichtfelds ist nicht auf die horizontale Dimension beschränkt. Dieselbe Methode kann auch zur vertikalen Erweiterung des Sichtfelds verwendet werden, um so eine echte sphärische Sicht zu erhalten. Die Lösung läuft auf einer Arm V8, 6-Kern-Edge-CPU mit 10 fps für vier Sensoren, die ein Sichtfeld von 275° liefern. Bei Verwendung von nur zwei Sensoren erhöht sich die Bildrate auf 30 fps.

Einer der Hauptvorteile dieses Ansatzes ist der massive Gewinn an Rechenleistung – mehr als eine Verdreifachung der Rechenleistung bei grundlegenden Berechnungen (siehe Tabelle 1).

Bild 8 und 9 zeigen einige Ergebnisse, die mit dieser Lösung erzielt wurden.

Wer schreibt:

Rajesh Mahapatra verfügt über mehr als 30 Jahre Berufserfahrung und arbeitet in der Gruppe Software und Sicherheit von Analog Devices in Bangalore.

Anil Sripadarao arbeitet in der Gruppe Software und Sicherheit von ADI in Bangalore.

Swastik Mahapatra ist Senior Machine Learning Engineer in der Gruppe Software und Sicherheit von ADI.

Literatur:

- „Analog Devices 3DToF ADTF31xx.“ GitHub, Inc.
- „Analog Devices 3DToF Floor Detector.“ GitHub, Inc.
- „Analog Devices 3DToF Image Stitching.“ GitHub, Inc.
- „Analog Devices 3DToF Safety Bubble Detector.“ GitHub, Inc.
- „Analog Devices 3D ToF Software Suite.“ GitHub, Inc.
- He, Yingshen, Ge Li, Yiting Shao, Jing Wang, Yueru Chen, and Shan Liu. „A Point Cloud Compression Framework via Spherical Projection.“ 2020 IEEE International Conference on Visual Communications and Image Processing, 2020.
- Industrial Vision Technology. Analog Devices, Inc.
- Topiwala, Anirudh. „Spherical Projection for Point Clouds.“ Towards Data Science, März 2020. ◀