

# IPS – ein System für eine mobile Datenerfassung in Innenräumen

Eugen Funk, Anko Börner, Ines Ernst, Jürgen Wohlfeil, Denis Griebach, Dirk Baumbach

Deutsches Zentrum für Luft- und Raumfahrt

Rutherfordstr. 2 12489 Berlin

{eugen.funk, anko.boerner, ines.ernst, juergen.wohlfeil, denis.griesbach, dirk.baumbach}@dlr.de

## Kurzfassung

*Vermessung und Inspektion von Innenräumen ist eine aufwendige Aufgabe. Die Fotos oder 3D Scans müssen immer räumlich zueinander referenziert sein, um die Konsistenz des Modells sicherzustellen. Erschwerend kommt hinzu, dass die Vermessung in oft zuvor unbekanntem oder nicht kartierten Bereichen stattfindet. Wir propagieren ein System zur gleichzeitigen Kartierung und Lokalisierung auf Basis von Bild- und Beschleunigungssensoren, sodass z.B. alle Fotos automatisch zueinander referenziert werden und sehr große konsistente 3D Modelle aufgebaut werden können.*

### 1. Einleitung

Die Erfassung von 3D-Geometrien mit Kameras rückt immer weiter in den Fokus von Vermessungstechnologien. Das Ziel Orte, Gebäude, Kunstobjekte oder kulturelles Erbe zu digitalisieren, geht einher mit der Digitalisierung unserer Welt und der Automatisierung vieler operationeller Prozesse. So wurde in erster Linie die Verwendung akkurater 3D Scanner wie dem Z-F-Imager 5016<sup>1</sup> von Zoller-Fröhlich favorisiert. Problematisch sind dabei vor allem zwei Aspekte: i) Die Messungen aus unterschiedlichen Positionen müssen zueinander „verortet“ werden, und ii) die Erfassung von Bereichen in engen Räumen ist nicht machbar.

Motiviert durch die beiden Schwierigkeiten haben wir am Institut für Optische Sensorsysteme am DLR ein handgetragenes Kamera-System entwickelt (s. Abbildung 3), welches i) 10 Bilder pro Sekunde aufnimmt, in Echtzeit verortet und ii) aufgrund der Größe und des



Abbildung 2: Handgetragenes Verortungssystem IPS.



Abbildung 1: a) Originalkamerabild, b) rekonstruiertes Tiefenbild.

Gewichts ohne Probleme in engen Umgebungen eingesetzt werden kann. So können 3D-Scanprozesse deutlich beschleunigt und eines Tages komplett automatisiert werden.

### 2. Systembeschreibung

Das integrierte Positionierungssystem (IPS) nutzt lediglich Bilder und Beschleunigungssensordaten zur Berechnung der Eigenbewegung (Abbildung 2). Es benötigt also kein absolutes Referenzierungssystem wie GPS, WLAN, RFID o.ä. Dadurch ist IPS in zuvor unbekanntem Umgebungen einsetzbar, ohne dass man diese vor der Vermessung mit Infrastruktur versehen muss.

Das System, zu sehen in Abbildung 3, besteht aus zwei grundlegenden Modulen: i) Sensoren und Navigation-Engine und ii) Algorithmen zur 3D Modellierung, welche auf Basis der Tiefenbilder (s. Abbildung 1) über die Zeit das komplette 3D Modell eines Gebäudes aufbauen können.

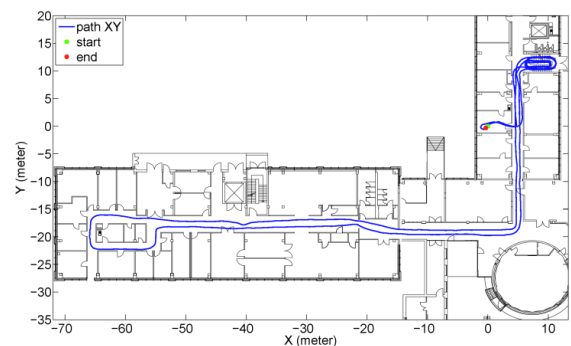


Abbildung 3: Die vom System ermittelte Trajektorie, getragen über 450m, eingezeichnet in den Grundriss des Gebäudes.

<sup>1</sup> Z-F-Imager 5016, <http://www.zf-laser.com>

## 2.1. Navigation-Engine

Durch die Sensordatenfusion gelingt es uns die Position und die Orientierung des Systems in Echtzeit zu bestimmen. Ursprünglich entwickelt für den Einsatz in der Raumfahrt, wurde sehr großer Wert auf die Genauigkeit der Lokalisierung gelegt: So erreicht IPS einen Driftfehler von ca 0.1% des Laufweges, was nur 10% des aktuellen Standes in der Forschung ist<sup>2</sup>. Mit dieser Genauigkeit ist es möglich, sehr große Entfernungen (z.B. 1000m) zurückzulegen, ohne einen anwendungskritischen Fehler (max 1m) zu riskieren. Falls eine noch höhere Genauigkeit in GPS-freien Umgebungen erforderlich sein sollte, wird die Verwendung von Markern empfohlen. Diese erlauben es, durch die Relokalisierung den Fehler deutlich zu verringern.

## 2.2. 3D-Modellierung

Die 3D Modellierung kann mit IPS auf zwei unterschiedliche Arten realisiert werden:

- i) Mithilfe der Stereobildpaare lassen sich bis 10 dichte Punktwolken pro Sekunde generieren und sukzessive einem 3D Modell hinzufügen. Dieses Prinzip wird im Abschnitt 2.2.1 beschrieben.
- ii) Wird eine hochauflösende Kamera mechanisch mit IPS verbunden, so lassen sich alle Fotos noch während der Aufnahme verorten. Diese Information kann als Vorwissen in ein Structure-From-Motion (SfM) System eingespeist (z.B. Agisoft) und weiter verfeinert werden. Desweiteren werden alle Bildbündel mithilfe eines Multi-View Stereo-Matching Algorithmus zu hochakkuraten 3D Punktwolken umgerechnet. Dieses Vorgehen ist im Abschnitt 2.2.2 beschrieben.

### 2.2.1 Sukzessive Echtzeitrekonstruktion

Die Bildpaare, die mit 10 Hz aufgenommen werden, lassen sich zur Berechnung der Tiefenbilder nutzen. Denn durch die Kenntniss des Abstandes der beiden Kameras zueinander wird die metrische Entfernung eines Objektes

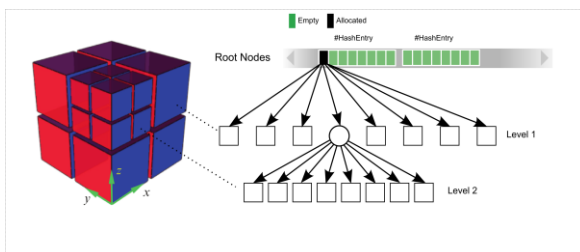


Abbildung 5: Strukturierung der 3D Punkte mit einem "Hashed-Octree".

in jedem Pixel berechnet. Wenn nun jeder Pixel eines Tiefenbildes zu einem 3D Punkt umgerechnet wird, erhält man eine dichte 3D Punktwolke pro Bildaufnahme.

Dadurch ergibt sich aber auch eine enorm hohe Datenrate, die nicht nur gespeichert, sondern auch visualisiert werden soll. Werden beispielsweise Bilder mit einer Auflösung von 640x480 Pixeln zu Tiefenbildern prozessiert, ergibt sich eine Datenrate von 307.200 3D-Punkten oder 18MB/Sekunde. Um die Datenflut verarbeiten zu können, verwenden wir sog. Hashed-Octrees<sup>3</sup> (s. Abbildung 5), womit sich unendlich große 3D-Punktwolken schnell und speicherschonend speichern und verarbeiten lassen.

Jedoch wird nicht jeder einzelne 3D-Punkt gespeichert. Diese werden zu kleinen Volumenelementen (Voxeln) gruppiert, sodass auf einem Volumen von z.B. 1cm<sup>3</sup> ein Messwert samt Orientierung aggregiert wird.

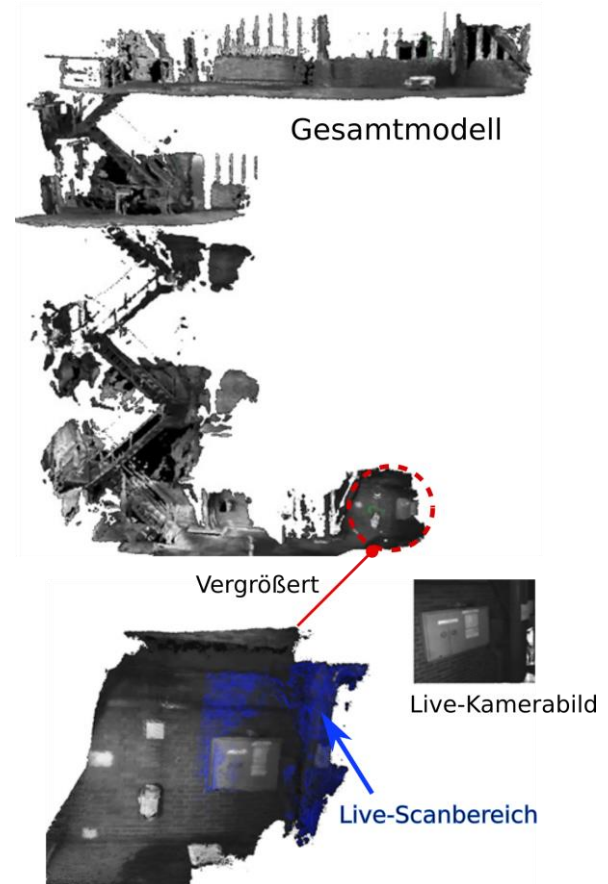


Abbildung 4: Die 3D Modellierung in Echtzeit ermöglicht es noch während der Vermessung den betrachteten Bereich zu selektieren.

<sup>2</sup> KITTI Benchmark: <http://bit.ly/2dku0Gc>, Stand 18.10.2016

<sup>3</sup> Funk E. and Börner A. (2016). Infinite 3D Modelling Volumes. In Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: VISAPP, ISBN 978-989-758-175-5, pages 246-253. DOI: 10.5220/0005722002460253

Das sukzessive Vorgehen ist von Vorteil, wenn noch während der 3D-Erfassung die Inspektorin oder der Inspekteur eine Rückmeldung vom System über die Scanqualität erhält (s. Abbildung 4). So können fehlerhafte Messkampagnen vermieden werden.

### 2.2.2 Hochgenaue Offlinerekonstruktion

Mit IPS ist man ebenfalls in der Lage hochakkurate 3D Modelle zu erzeugen. Mithilfe einer mechanisch verbundenen Kamera ist es möglich, sehr hochauflösende Fotos mit einer Objektauflösung deutlich unter 1mm zu verorten. Diese Position und Orientierung zu jedem Bild wird als Startwert für eine SfM Software verwendet<sup>4</sup>. Bei Bedarf lassen sich auch Informationen aus den Bodenkontrollpunkten integrieren und so die Drift der Verortung in größeren Umgebungen beseitigen. Das Ergebnis nach SfM sind verfeinerte Bündel von Bildern mit dazugehörigen Positionen und Orientierungen in metrischen Koordinaten. Diese werden nachfolgend von „Multi-View-Stereo-Matching“ zu dichten 3D-Punkten verarbeitet. Die Prozesskette ist in Abbildung 8 dargestellt. Abbildung 7 zeigt die Inspektionsergebnisse eines Schachts, wobei IPS gekoppelt an eine hochauflösende Kamera (Abbildung 6) an einem Drahtseil heruntergelassen wurde.

Der besondere Vorteil der vorgestellten Offlinemethode gegenüber Standard-SfM Verfahren ist die drastisch verringerte Rechenzeit: Die IPS-Trajektorie erlaubt es, das Bildmatching von quadratischer auf lineare Komplexität zu reduzieren. So gelingt die Prozessierung eines Datensatzes mit 100.000 Bildern innerhalb weniger Tage, was mit Standard-SfM einige Wochen dauern würde.

### 3. Genauigkeiten

Die Genauigkeit des Systems lässt sich mit zwei voneinander unabhängigen Größen beschreiben: i) Die Drift der Trajektorie beträgt ca 0.1% des Laufweges. Dieser Fehler lässt sich weiter auf nahezu Null mithilfe von referenzierten Markern reduzieren. ii) Die Genauigkeit der 3D-Rekonstruktion kann beliebig hoch, abhängig von der Optik und den Objektabständen, gestaltet werden. Die Qualität der echtzeitfähigen Rekonstruktion des „Vorschaumodells“ hingegen

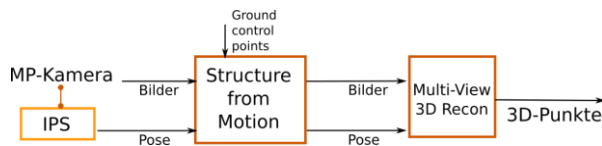


Abbildung 8: Prozessstruktur zur Generierung von hochakkuraten 3D Punktwolken aus IPS-verorteten Bildern.

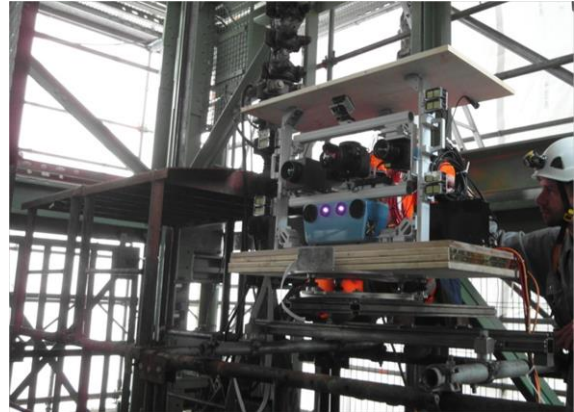


Abbildung 6: IPS (blau) gekoppelt mit zwei hoch auflösenden DSLR Kameras.

beschränkt sich auf wenige Zentimeter und hängt stark vom Nutzer ab (Bewegungsunschärfe, Helligkeit, etc.).

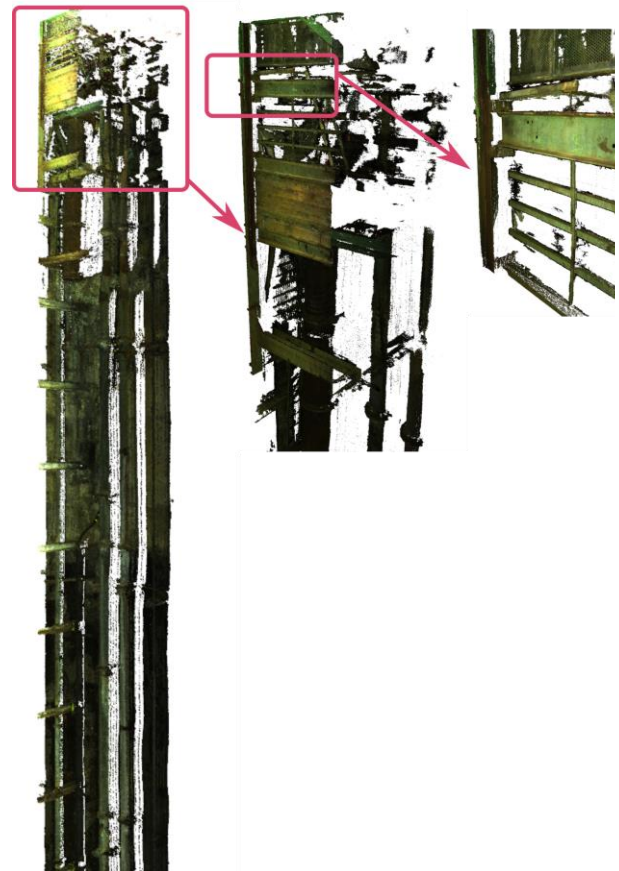


Abbildung 7: Rekonstruktion eines Schachts mithilfe der Offlineprozessierung.

<sup>4</sup> Z.B. Photoscan von Agisoft

#### 4. Ausblick

Das System wird in den kommenden Monaten weiter evaluiert, sodass für unterschiedliche Szenarien generalisierte, quantitative Ergebnisse generiert werden. Aktuell befinden sich die ersten Prototypen des Systems im Aufbau, die ab 2017 von der Firma DMT in Kleinserie produziert werden. So erreichen wir, dass die Technologie von allen Vermessungsbüros genutzt werden kann.

#### 5. Zusammenfassung

Das vorgestellte IPS System liefert eine hochpräzise Trajektorie, basierend auf Sensordatenfusion von Bildern und Beschleunigungssensormessungen. Mithilfe der Stereobilder wird es ermöglicht, in Echtzeit die 3D-Geometrie eines Objektes zu rekonstruieren. Weiterhin wird die in Echtzeit bestimmte Trajektorie zur effizienten hochakkuraten und nachträglichen 3D-Rekonstruktion genutzt. So werden die hierfür notwendigen Rechenzeiten um ein Vielfaches reduziert.

IPS kann in vielen Bereichen zur Anwendung kommen, bei denen die räumliche Referenzierung von Messdaten und eine darauf basierende Modellbildung gefordert wird und externe Systeme, wie GPS, nicht zur Verfügung stehen. Das gilt auch für die Erfassung von Kunst- und Kulturobjekten – seien es Ausgrabungsstätten, Gebäude oder große Ausstellungsgegenstände.