

# **3D UMGEBUNGSREKONSTRUKTION IN ECHTZEIT – EINE AUTOMOTIVE ANWENDUNG LERNT DAS FLIEGEN**

Markus Prochaska, Georg Leugner, Dr. Emanuel Brämer,  
ESG Elektroniksystem- und Logistik GmbH, 82256 Fürstenfeldbruck, Deutschland

## **Zusammenfassung**

Im Rahmen eines internen Technologie- und Innovationsprogramms ist es der ESG gelungen, Schlüsseltechnologien aus dem Bereich Automotive, wie z.B. Simultaneous Localisation and Mapping (SLAM), exemplarisch auf aktuelle Anwendungen in der Luftfahrt zu übertragen. Als Grundlage dieses Vorhabens diente ein Algorithmus für monokulares „Structure from Motion“ (SfM), der für ein Fahrerassistenzsystem aus dem Bereich Augmented Reality entwickelt wurde. Dieser Algorithmus wurde dahingehend angepasst, dass er nun für Luftfahrzeuge zur Verbesserung der Positionsbestimmung einsetzbar ist.

Im Bereich der Automotive Navigation werden bei der ESG verschiedene Projekte durchgeführt, die eine kontaktanaloge Anzeige von Navigationshinweisen im Sichtfeld des Fahrers – z.B. mittels eines Head-Up-Displays – zum Ziel haben. Um die Navigationshinweise stabil zu platzieren, muss die Position und Orientierung des Fahrzeugs bzgl. seiner Umgebung zu jedem Zeitpunkt mit höchster Präzision bekannt sein, was mit Hilfe der SfM-Technologie erreicht werden kann.

Die verwendeten Algorithmen wurden mit wenigen Modifikationen auf die Anwendung an Bord von Remotely Piloted Aircraft Systems (RPAS) übertragen und dienen hier dazu, die Bestimmung der absoluten Position der fliegenden Plattformen zu stützen. Dazu muss die errechnete hochpräzise relative Position mit den Positionsinformationen aus anderen Sensoren - insbesondere GPS - fusioniert werden. Die Sensorfusion von GPS und relativen Positionsinformationen wurde in Flugversuchen untersucht. Hierbei konnte die Gangbarkeit des Ansatzes SfM zur Unterstützung der Positionsbestimmung in RPAS nachgewiesen werden.

## 1. EINLEITUNG

Die ESG Elektroniksystem-und Logistik-GmbH ist in vielen verschiedenen Technologiezweigen tätig und hat dadurch die Möglichkeit das erreichte Wissen zwischen den verschiedenen Bereichen zu transferieren. Gerade zwischen der Automotive- und Luftfahrt-Domäne ist es in vielen Fällen sinnvoll vorhandene Technologien durch geringen Aufwand anzupassen. Durch die langjährige Erfahrung im Bereich Fahrerassistenzsysteme für mehrere Automobilhersteller hat der ESG Automotive Bereich unter anderem komplexe kamerabasierte Verfahren und Algorithmen entwickelt.

In einem internen Technologieprojekt sollte die für Autos entwickelte Structure from Motion (SfM) Technologie für die Anwendung im Bereich unbemannter Fluggeräte untersucht werden. Das im ESG Automotive Bereich entwickelte SfM-Verfahren wurde erfolgreich in mehreren Demonstratoren im Umfeld der Fahrerassistenzsysteme eingesetzt. Als Beispiele sind hier die Online Calibration, Augmented Reality und Surround View zu nennen.

Um die im Automotive Bereich vorhandene Technologie im Avionik Umfeld zu testen, wurden zwei Multicopter der Firma Ascending Technologies mit einer Bluefox Kamera als Informationsquelle ausgestattet. Da die Anzahl der Freiheitsgrade in der Luftfahrt die des Fahrzeugs am Boden übersteigt, musste der existierende Algorithmus den Begebenheiten angepasst werden.

Die Echtzeitfähigkeit der verwendeten SfM-Komponente und die verhältnismäßig geringen Kosten der Hardwarekomponenten ermöglichen im Automobil sowie im Luftfahrtsektor einen Einsatz in unterschiedlichsten Anwendungen.

Mit Hilfe der übertragenen Technologie soll die Positionsbestimmung des Multicopters durch eine Fusionierung der verfügbaren Daten (GPS, Inertialsensoren, SfM) verbessert werden. Durch diese Erweiterung kann die Bandbreite der Anwendungen von unbemannten Plattformen weiter erhöht und beispielsweise auch ein Einsatz in urbanen Umgebungen mit vielen Hindernissen bewerkstelligt werden.

Das erarbeitete Wissen, das mit Multicoptern im Kleinen gezeigt wurde, soll in Zukunft auch auf großen Plattformen eingesetzt und weiterentwickelt werden. Hierfür soll der Unbemannte Missionsausrüstungsträger (UMAT) der Firma ESG genutzt werden, der mit verschiedenen Nutzlasten und Autonomiegraden eingesetzt werden kann.

## 2. GRUNDLAGEN

Es gibt in der heutigen Zeit eine große Anzahl von Projekten in denen Simultaneous Localisation and Mapping (SLAM) sowie Structure from Motion Techniken eingesetzt werden. Gerade im Bereich der Robotik sind viele Open-Source und kommerzielle Lösungen zu finden, da erst durch die simultane Lokalisierung und Kartierung die autonome Navigation in einer unbekanntem Umgebung bewerkstelligt werden kann. Als

Informationsquellen für SLAM werden oft teure Techniken wie Laserscanner, Radar oder Sonar genutzt. Um die Größe und Kosten der erforderlichen Geräte möglichst klein zu halten wurde in diesem Projekt als Sensor eine Kamera gewählt. Durch dieses monokulare Kamerasystem soll die Genauigkeit der für die Navigation der unbemannten Plattform notwendigen Positions- und Geschwindigkeitsdaten verbessert werden.

### 2.1. Kamera, GPS und IMU

Um die Bewegungsvorgänge einer unbemannten Plattform in allen sechs Freiheitsgraden zu messen, stehen mehrere Informationsquellen zur Verfügung. Im hier vorgestellten Projekt wurden die Inertialsensoren der verwendeten Multicopter, GPS und eine Kamera benutzt.

#### 2.1.1. Kamera

Bei Verwendung eines Monokamerasystems wird die 3D-Welt auf 2D-Bilder, also eine Ebene projiziert. Bei diesem Vorgang gehen die Tiefeninformationen verloren, wodurch auch die Größe von Objekten nicht gemessen werden kann. Das heißt es ist durch das projizierte 2D-Bild nicht möglich die Strecke zwischen einem Pixel auf dem Bild und der Kamera zu erkennen. Erst durch die Bewegung der Kamera durch den Raum und den dadurch erzeugten optischen Fluss, lassen sich diese Informationen wieder gewinnen.

Bei der Auswahl der Kamera Hardware sollte auf eine hohe Auflösung geachtet werden um kleinste Strukturen im Bild wiedererkennen zu können. Eine ausreichend hohe Bildfrequenz ist nötig um schnelle Bewegungsabläufe, die gerade bei wendigen Plattformen wie Multicoptern vorkommen, zu erfassen.

Ein weiteres wichtiges Feature für den Anwendungsfall, dass sich die Kamera an Bord eines Multicopters befindet, ist die Ausstattung mit einem globalen Shutter. Das bedeutet, dass ein Bild nicht zeilenweise, sondern im Ganzen aufgebaut wird um somit die sogenannten Rolling Shutter Effekte (BILD 1) zu vermeiden.



BILD 1. Rolling Shutter Effekt

#### 2.1.2. GPS

Auf den verwendeten Multicoptern ist ein Global Positioning System (GPS) Empfänger zur Positionsbestimmung vorhanden. Auch wenn die Verwendung der GPS Position im Bereich Luftfahrt und

Automotive weit verbreitet ist, sollte die erhaltene Position wegen der vorhandenen Ungenauigkeit nur als unterstützende Funktion benutzt werden. Durch die hier beschriebene Fusion mit anderen Sensoren kann die Genauigkeit und Ausfallwahrscheinlichkeit der Position verbessert werden.

### 2.1.3. IMU

Zur Messung von Beschleunigungen und Drehraten aller sechs Freiheiten einer Plattform dient die Kombination der auf dem System angebrachten Inertialsensoren. Diese Kombination unterschiedlicher Messeinheiten wird Inertial Measurement Unit (IMU) genannt. Um alle sechs Freiheiten zu erfassen werden drei orthogonal angeordnete Beschleunigungssensoren und drei orthogonal angeordnete Drehratensensoren (Gyroskope) benötigt. Die im Projekt verwendeten Multicopter sind zur Steigerung der Genauigkeit zusätzlich noch mit einem Kompass ausgerüstet.

Die Positionsbestimmung alleine durch die Inertialsensoren ist, wenn überhaupt, nur für wenige Sekunden möglich, da mit fortschreitender Messdauer der Positionsfehler quadratisch ansteigt. Erst durch die Fusion mit GPS ist eine hohe Genauigkeit bei der Positionsbestimmung möglich.

## 2.2. Einsatz im Automotive Bereich

Kameras werden im Automotive Bereich zur Bildverarbeitung auf Grund Ihrer Vorteile im sichtbaren Bereich seit längerer Zeit eingesetzt. Die Systeme sind kostengünstig, es gibt keine elektromagnetische Belastung und sie ermöglichen die bildhafte Interpretation der Szene.

Trotz der genannten Vorteile gibt es auch Schwierigkeiten, die beim Einsatz beachtet werden müssen, wie zum Beispiel die Größe der zu verarbeitenden Datenmenge. Bei Fahrzeugen können zur Berechnung der Eigenbewegung des Fahrzeugs bei monokularen Kamerasystemen die im Auto verfügbaren Daten wie Geschwindigkeit und Drehzahlen herangezogen werden.

In heutigen Fahrzeugen angewendete Techniken im Bereich Bildverarbeitung steigern die Sicherheit enorm. Als Beispiel können hier die Kollisionswarnung beim Einparken und Rückwärtsfahren, die Spurerkennung, die Fernlichtautomatik, Fußgängererkennung und die Auffahrwarnung genannt werden.

Die immer steigende Anzahl von Anwendungsmöglichkeiten im Bereich Fahrerassistenzsysteme und das in der ESG vorhandene Know-how lassen auf eine sinnvolle Übertragbarkeit der Technik in die Luftfahrt schließen.

## 2.3. Möglichkeiten in der Luftfahrt

Im Bereich Luftfahrt ist der Einsatz von kamerabasierter Sensorik genauso sinnvoll wie im Automobilsektor. Gerade bei unbemannten Plattformen ist das Erkennen der sichtbaren Umgebung ein wichtiger Sicherheitsfaktor.

Viele heute benutzte Plattformen verlassen sich auf die korrekte Positionsbestimmung durch GPS. Doch gerade in urbanen Umgebungen, in Bauwerken oder bei Störung des GPS Signals durch Dritte, darf GPS nicht die einzige Quelle der Positionsbestimmung sein. Erst durch die Fusion verschiedener Positionsquellen wie SLAM und GPS wird die Genauigkeit und die Zuverlässigkeit der ermittelten Daten größer.

Als weitere Anwendungsmöglichkeiten im Bereich Luftfahrt können die Aufklärung von Landezonen, oder die Hindernisvermeidung genannt werden. Diese spielt gerade bei unbemannten Plattformen eine große Rolle, da dadurch die Kollision mit anderen Luftfahrzeugen, dem Gelände oder Personen verhindert werden kann.

## 3. ESG SLAM KOMPONENTE

Bei der von der ESG entwickelten SfM-Technologie handelt es sich um die simultane visuelle Bewegungsschätzung und die 3D-Rekonstruktion der Umgebung mittels einer Videokamera. Eine in C++ prototypisch umgesetzte Softwarekomponente nimmt einen Videostrom und die intrinsische Kamerakalibrierung (Brennweite, Linsenverzerrung etc.) als Eingabe entgegen und errechnet in Echtzeit zu jedem Videobild eine entsprechende Pose (Translation + Orientierung) sowie eine dreidimensionale Punktwolke.

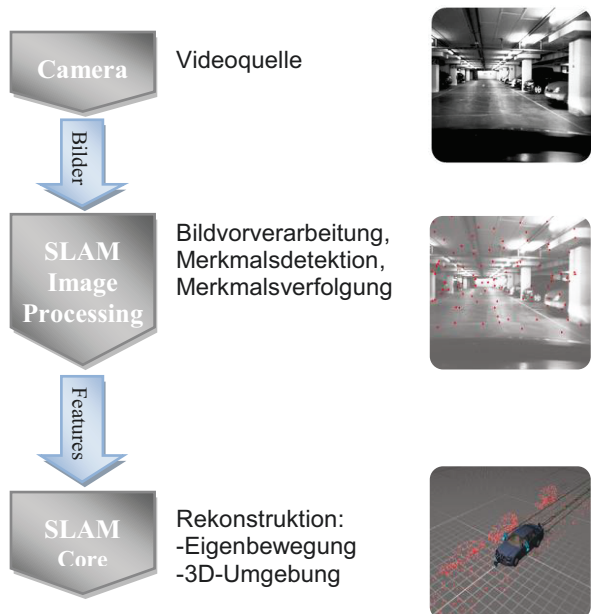
Wichtig anzumerken ist dabei, dass die SfM-Komponente sich vollkommen agnostisch gegenüber der Bewegungsart verhält, solange eine (translatorische) Bewegung überhaupt stattfindet.

### 3.1. Design

Während im Automobilbereich die komplette Sensordatenverarbeitung (noch) im Fahrzeug selbst stattfindet, ist beim Einsatz unbemannter Fluggeräte (UAV) die Möglichkeit der verteilten Datenverarbeitung zu berücksichtigen. Dabei wird der Rechenaufwand zwischen dem Bordrechner des UAVs und einer Bodenstation geteilt. Eine solche Aufteilung soll für die flexible Ressourcenverteilung sorgen, wodurch der variierenden Leistungsfähigkeit der Bordrechner, sowie der Kommunikationskanäle Rechnung getragen werden kann.

Die zu erstellende Architektur sollte daher sowohl einen „self-contained“ Einsatz am Bord eines UAVs ermöglichen, als auch die Verschiebung der Rechenlast auf eine Bodenstation unterstützen.

Die Software-Pipeline wurde im Rahmen des Projekts in folgende Bausteine aufgeteilt, die untereinander per Netzwerk kommunizieren:



Alle drei Bausteine können ihren Dienst auf einem einzigen Rechner verrichten, oder aber auf unterschiedlichen Rechnern installiert werden. So ist es z.B. möglich, den Videostrom einer am UAV angebrachten Kamera drahtlos zur Bodenstation zu transferieren, wo alle weiteren Schritte der SfM-Auswertung stattfinden. Bei geringer Bandbreite des Kommunikationskanals ist es andererseits denkbar, die Bildvorverarbeitung mit anschließender Merkmalsverfolgung auf dem UAV zu verrichten, um nur die 2D-Merkmale zur Bodenstation zu übermitteln.

Um die Ergebnisse der Eigenbewegungsschätzung und der 3D-Rekonstruktion intuitiv erfassbar zu machen, wurde die vorhandene Open-Source Visualisierungsumgebung „rviz“ um ein proprietäres Plugin erweitert (siehe Bild 2). Rviz kam gleichermaßen für die Darstellung der Bewegungstrajektorien wie der 3D-Punktewolken zum Einsatz.

### 3.2. Bewegungsschätzung

Die Herausforderung bestand in der Verknüpfung des rekursiven und inkrementellen SfM-Verfahrens zur visuellen Bewegungsschätzung mit den absoluten Daten des GPS. Hierbei sollten die Stärken der beiden Positionierungsverfahren kombiniert sowie die Schwächen neutralisiert werden.

In diesem Zusammenhang ist die hohe lokale Genauigkeit sowie die - gegenüber GPS - hohe Messfrequenz (bis zu 30 Hz) der rekursiven visuellen Bewegungsschätzung zu nennen. Auf der anderen Seite leidet die rekursive visuelle Bewegungsschätzung - wie alle inkrementellen Verfahren - unter der unbeschränkten Fehlerakkumulation. Demgegenüber arbeitet ein GPS-Empfänger üblicherweise mit deutlich niedrigeren Messfrequenzen (typischerweise 1 - 5 Hz) und bietet bei einem konstanten globalen Fehler eine unzureichende lokale Positionierungsgenauigkeit von mehreren Metern.

Hierbei galt es, zwei Probleme zu lösen. Zum einen sollte die räumliche Transformation zwischen den Weltkoordinatensystemen der beiden Informationsquellen automatisch ermittelt, und zum anderen eine möglichst optimale Pose (Position und Orientierung) bezüglich des GPS-Koordinatensystems errechnet werden.

### 3.3. Visuelle Bewegungsschätzung mittels SfM

Bei der visuellen Bewegungsschätzung wurde ein am UAV befestigtes monokulares Kamerasystem eingesetzt, dessen Videostrom der SfM-Komponente zugeführt wurde. Die intrinsischen Parameter der Videokamera wurden durch ein Kalibrierungsverfahren gewonnen. Die SfM-Komponente errechnete in Echtzeit aus dem Videostrom die Bewegung der Kamera in 5 Freiheitsgraden (3 Winkel der Orientierung + 3 Bewegungsparameter - Skalierung der Szene).

In BILD 2 sieht man einen Ausschnitt der Rviz Visualisierungsumgebung. Dargestellt werden die von SfM berechnete Trajektorie (rot), die GPS-Daten (blau) und die 3D-Punktewolke (hellrot). Die Sequenz entstand durch das Tragen des UAVs. Man beachte die Abbildung der Schritte in der SfM-Kurve sowie die Ungenauigkeit des GPS-Tracks.

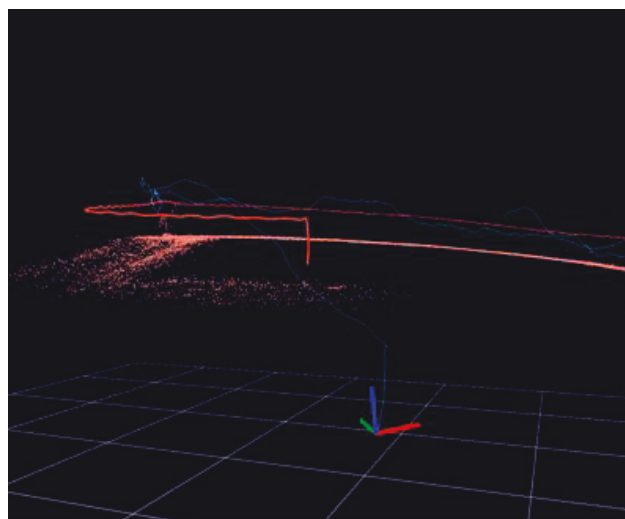


BILD 2. Rviz Visualisierungsumgebung

### 3.4. Sensorfusion mit GPS

Die Ermittlung der Transformation zwischen dem SfM-Weltkoordinatensystem und dem GPS-Weltkoordinatensystem erfolgte durch einen erweiterten Informationsfilters. Hierbei dienten die synchron aufgenommenen Datenpaare (Positionen) aus SfM und GPS als Eingabedaten, wobei die Transformation zwischen den Koordinatensystemen von SfM und GPS laufend aktualisiert wurde. Um die Aussagekraft der GPS-Daten zu verbessern, wurden nur Daten mit einem hinreichend großen Signal-Rausch Abstand akzeptiert.

### 3.5. 3D Rekonstruktion

Wie der Name suggeriert, ist ein Ergebnis der SfM-Technologie die dreidimensionale Rekonstruktion der Umgebung. Eine mögliche Anwendung im

Zusammenhang mit unbemannten Fluggeräten besteht in der Exploration und der Kartierung eines unbekanntes Gebiets. Die in 3D erfasste Landschaft kann wiederum zur Verbesserung der Navigation herangezogen werden, insbesondere im Kontext der Kollaboration mehrerer UAVs.

Die Ausrichtung der rekonstruierten 3D-Umgebung im GPS-Koordinatensystem erfolgt mittels der räumlichen SfM-GPS Transformation, die dabei im Rahmen der Sensorfusion ermittelt wurde.

Um aus erkannten 3D-Punkten kolorierte Oberflächen zu erstellen, werden die in BILD 3 dargestellten Arbeitsschritte durchlaufen.

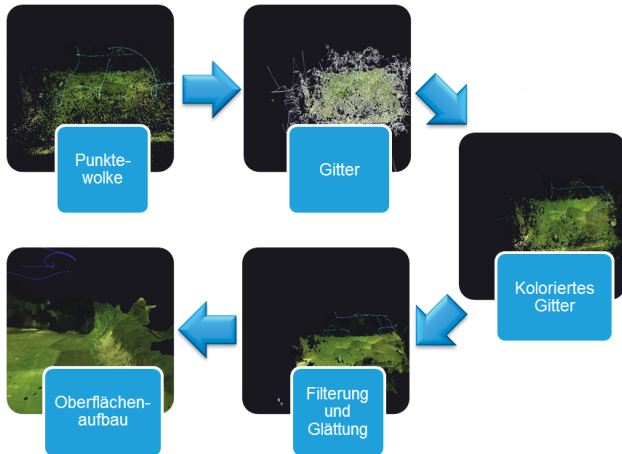


BILD 3. Von Punkten zu Oberflächen

Als erstes muss die 3D-Punktewolke erstellt werden, die sich aus den erkannten Features zusammensetzt. Diese Features sind Bildpunkte in der realen Welt, die durch spezielle Merkmale leicht wiedererkannt werden können. Die räumliche Wahrnehmung und Tiefenbestimmung wird durch die Verfolgung dieser Features über mehrere Bildfolgen hinweg ermöglicht. In BILD 4 werden die erkannten Features der aufgenommenen Szene dargestellt.

Die erkannten 3D-Punkte werden zur Veranschaulichung in diesem Bild in verschiedenen Farben dargestellt. Dies soll den Umgang der SLAM Komponente mit Features verdeutlichen. Es ergeben sich folgende Bedeutungen für die verschiedenen Färbungen der Punkte:

- Blau: In Blau werden die neu erkannten Features dargestellt.
- Rot: Die Farbe Rot zeigt die nicht wieder erkannten Features, die verworfen werden.
- Rosa: In Rosa werden die vom Algorithmus wiedererkannten Punkte dargestellt.

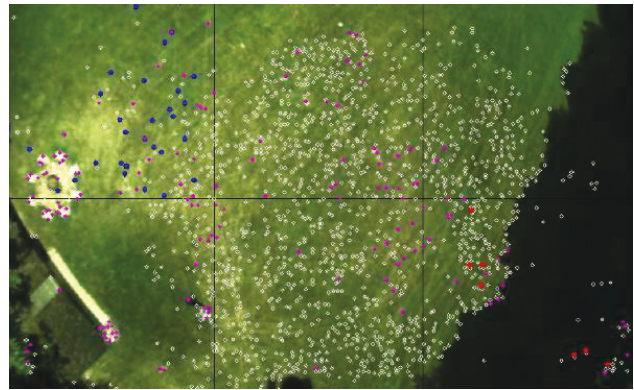


BILD 4. Bildfeatures

Der nächste Schritt ist das Kolorieren der 3D-Punktewolken. In der ersten Ausbaustufe wurde jedem 3D-Punkt in einer 3D-Punktewolke eine Farbe zugeordnet. Diese ergab sich aus der Auswertung der Videobilder, in denen der jeweilige 3D-Punkt zu sehen war. Um sicherzustellen, dass für jeden 3D-Punkt eine Farbe ermittelt werden kann, wurde eine Reihe von Videobildern – sog. Keyframes – zwischengespeichert (BILD 5).

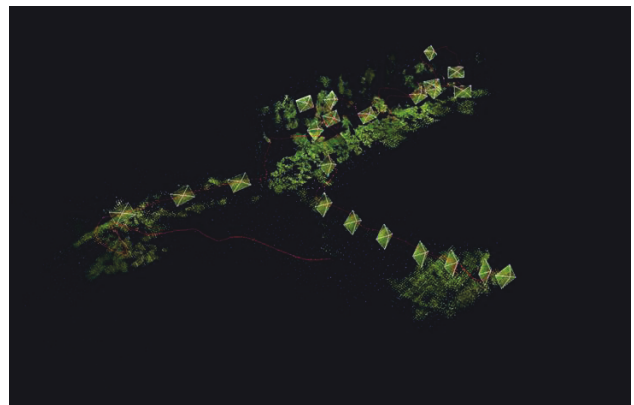


BILD 5. Keyframes und die kolorierte 3D-Punktewolke

Im nächsten Schritt wurden die errechneten 3D-Punkte zu einem zusammenhängenden dreidimensionalen Netz verknüpft, was unter der Zuhilfenahme der frei verfügbaren PCL Bibliothek geschah, wobei die Farbe zwischen den 3D-Punkten interpoliert wurde.

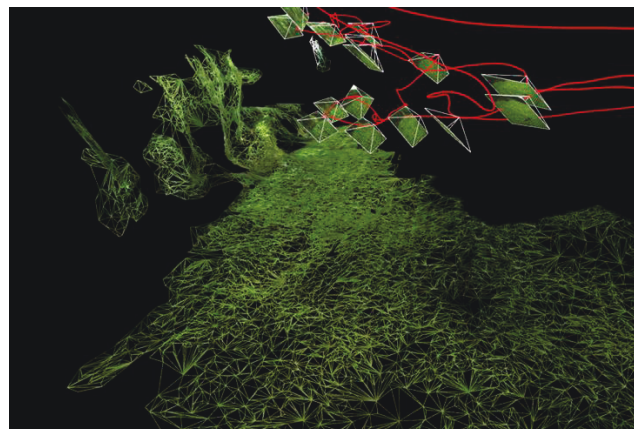


BILD 6. Das aus der 3D-Punktewolke erzeugte Netz mit interpolierten Farben

Anschließend wurde eine weitere Darstellungsform geschaffen, bei der nicht die eigentliche 3D-Punktwolke koloriert wird, sondern die Videobilder auf eine Fläche projiziert werden, die durch die Analyse der 3D-Punktwolke ermittelt wird (BILD 7).

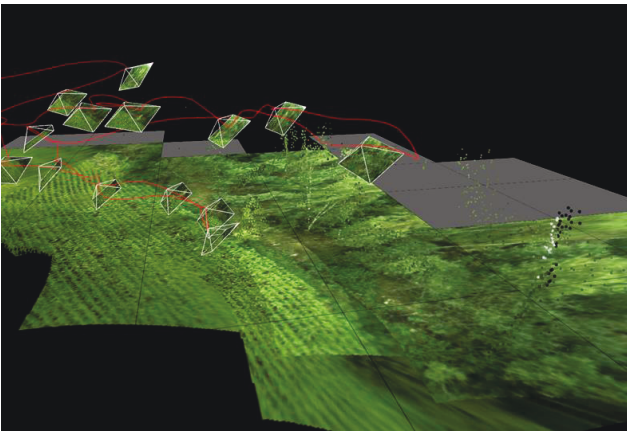


BILD 7. Projektion der Videobilder auf eine im GPS-Koordinatensystem ausgerichtete Bodenebene

### 3.6. Einschränkungen

Bei ungünstigen Bedingungen wie der dominanten Rotationskomponente der Bewegung, der Bewegungsunschärfe oder Ähnlichem, bricht die visuelle Schätzung ab. Dieses Verhalten kann mehrfach hintereinander auftreten, sodass die Länge der verfügbaren Bewegungsschätzung nicht ausreicht, um sicher eine SfM-GPS Transformation zu bestimmen. Als Folge davon kann die 3D-Punktwolke nicht im GPS-Koordinatensystem ausgerichtet werden, was zu mehreren im Raum verteilten 3D-Punktwolken führt.

### 3.7. Ausblick

Um die Ergebnisse, die mit Hilfe der Multicopter erarbeitet wurden, auf größeren Plattformen zu testen, soll der Algorithmus mit einem großen UAV getestet werden. Hierfür wird eine Portierung auf den Unbemannten Missionsausrüstungsträger (UMAT) der ESG durchgeführt.

Weiterhin soll eine funktionale Weiterentwicklung im Bereich semantische Analyse der 3D Rekonstruktion durchgeführt werden. Auch der Einbezug weiterer Quellen wie Laser Scanner oder Infrarot Sensoren zur Sensorfusion sollen untersucht werden.