

Anwendung von Motion-Tracking zur Erstellung von 3D-Szenen und Vermessung darin enthaltener Objekte

Sven Becker, Tim Hanjo Fritzsich, Dirk Labudde

Kontakt: Sven Becker, Hochschule Mittweida, becker1@hs-mittweida.de

Zusammenfassung

Die Digitalisierung forensischer Ereignis- und Tatorte hat mittlerweile Einzug in die tägliche Arbeit von Strafverfolgungsbehörden gefunden. Basis für die Erstellung sogenannter 3D-Szenen sind, neben Laserscannern, meist Videoaufnahmen, die vom Boden oder aus der Luft heraus erhoben werden. Die Digitalisierung erfolgt nicht nur zu Dokumentationszwecken, sondern auch um Objekte in den erstellten digitalen Szenen zu vermessen. So robust die Vielzahl an Algorithmen, wie Structure-from-Motion, für die 3D-Erstellung sind, so anfällig sind diese z.B. bei Videomaterial mit minderwertiger Ausleuchtung. Jedoch weisen eine Vielzahl forensisch sowie rechtsmedizinischer Fälle, in denen Täter oder Opfer videographisch erfasst wurden, ebensolches Material als Analysegrundlage auf. Aus diesem Grund erfolgt in diesem Beitrag die Vorstellung eines Workflows zur Erstellung von 3D-Szenen auf Grundlage von Motion-Tracking sowie von Möglichkeiten zur Vermessung von Objekten in 3D-Szenen. Vorteil des Workflows ist, dass lediglich ein zur Verfügung stehendes Video für eine digitale Szenenerstellung ausreicht und keine Vor-Ort-Aufnahmen des Ereignis- oder Tatortes notwendig sind. Der entwickelte Workflow kann unterstützend für Ermittlungen und Analysen eingesetzt werden und bringt einen enormen Mehrwert im Bereich der forensischen 3D-Rekonstruktion und -Vermessung.

Keywords: Motion Tracking, 3D-Rekonstruktion, Objektvermessung.

1 Einführung

Die 3D-Rekonstruktion und Analyse von Objekten auf Grundlage von Bild- und Videomaterial stellt mittlerweile ein Standardvorgehen für Ermittlungsbehörden dar. Algorithmen wie Structure-from-Motion erstellen innerhalb kürzester Zeit komplexe 3D-Szenen aus zweidimensionalem Bild- und Videomaterial. Vorteil dieser Algorithmen ist, dass sie vollautomatisch funktionieren. Ziel der Überführung von 2D-Daten in 3D-Modelle ist zumeist die Vermessung von in den Szenen enthaltenen Objekten, bspw. von Spuren oder Personen. Die Vermessung solcher Objekte ist in verschiedenen wissenschaftlichen und industriellen Anwendungen wie der Archäologie, der Robotik, der medizinischen Bildgebung und der virtuellen Realität von großer Bedeutung. Meist sind jedoch digitale Szenenerstellungen und Vermessungen bei z.B. schlecht ausgeleuchtetem Videomaterial nicht möglich. Algorithmen wie Structure-from-Motion kommen hier an ihre Grenzen. Jedoch besteht oft vor allem bei solchem Analysematerial die Herausforderung für Ermittlungsbehörden darin, präzise und effektive Methoden zur Objektvermes-

sung zur Verfügung zu haben. Vor allem, wenn diese von entscheidender Bedeutung sind, um Erkenntnisse über die Abmessungen, Proportionen und strukturellen Merkmale von zum Beispiel möglichen Verletzungen eines Opfers oder über die Körpergröße eines Täters zu gewinnen. Dieses Wissen kann helfen, den zeitlichen Rahmen eines Tatgeschehens einzugrenzen. Bestehende Methoden zur Vermessung von Objekten direkt im Bild- und Videomaterial haben jedoch ebenfalls ihre Grenzen. Die Genauigkeit und Effizienz dieser Methoden reichen oft nicht aus, um präzise Messungen in dem Bild- und Videomaterial durchzuführen. Daher besteht die Notwendigkeit, neue Ansätze zu entwickeln, die zunächst aus dem Bild- und Videomaterial eine zuverlässige und genaue 3D-Rekonstruktion der Szene und anschließend eine Vermessung von darin enthaltenen Objekten ermöglichen. In diesem Zusammenhang konzentriert sich die vorliegende wissenschaftliche Arbeit auf die Untersuchung von Techniken, wie Motion-Tracking zur 3D-Rekonstruktion solcher Szenen. An einem Beispiel erfolgt die Überprüfung der Genauigkeit von Vermessungen von Objekten, die in diesen Szenen enthalten sind. Die Ergebnisse dieser Arbeit werden einen wichtigen Beitrag zur Forschung leisten und auch in anderen Bereichen anwendbar sein, in denen die 3D-Erstellung von Szenen auf Grundlage schlechten Videomaterials und die Vermessung von Objekten von entscheidender Bedeutung sind. Die vorliegende Arbeit gibt im zweiten Kapitel einen Überblick über den aktuellen Stand der Forschung unter Einbezug ausgewählter Literatur. Aufbauend darauf erfolgt im dritten Kapitel die Einordnung der Arbeit in den Fach- und Forschungsstand. Der generelle Aufbau des entwickelten Workflows wird in Kapitel vier beschrieben. Kapitel fünf umfasst die Beschreibung des Workflows an einem Beispiel. Die letzten beiden Kapitel geben eine Zusammenfassung und einen Ausblick für weiterführende Analysen und Forschungen sowie eine Einordnung der Bedeutung und des Potentials des Workflows für die Ermittlungsarbeit wieder.

2 Stand der Forschung

Im Folgenden werden einige wissenschaftliche Beiträge aus dem Fachbereich der 3D-Rekonstruktion und Vermessung von Objekten aufgeführt, die noch heute ihre Gültigkeit besitzen und die Grundlage für die angewandten Vorgehensweisen gebildet haben.

Buck (2007) diskutiert den Einsatz von 3D-Technologien und -Methoden in der forensischen Untersuchung von Verkehrsunfällen, insbesondere zur Bestimmung der Aufprallsituation und Korrelation von Verletzungen mit den verursachenden Objekten. Die verwendeten Methoden umfassen 3D-Oberflächenscans für äußere Körperbefunde,

post-mortem MSCT und MRT-Scans für innere Befunde sowie computerunterstützte 3D-Vergleiche von Verletzungen und Fahrzeugbeschädigungen. Der Artikel gibt Beispiele für die Vorteile dieser Techniken bei der Analyse des Unfallgeschehens. Die Studie von Campana (2016) untersucht die Umsetzbarkeit der Dokumentation von strukturierten Verletzungen mit 3D-Technologie und Farbfotografie, ohne aufwendige 3D-Oberflächendokumentationsmethoden zu verwenden. Es wurde gezeigt, dass die Integration von Fotografien in CT/MRT-Datensätze für die 3D-Dokumentation von strukturierten Verletzungen geeignet ist. Allerdings ist dies kein Ersatz für Photogrammetrie und Oberflächenscans, speziell wenn präzise Daten für die vergleichende Analyse von hoch detaillierten Verletzungen mit dem verursachenden Instrument erforderlich sind. Bornik et al. (2018) beschreiben, wie immer häufiger 3D-Scanner und medizinische Bildgebung wie CT und MRT in der forensischen Arbeit verwendet werden. Er zeigt, wie digitale Aufzeichnungen dazu dienen können, Verletzungen und Spuren an Tatorten umfassend und nicht-invasiv zu dokumentieren. Er geht darauf ein, wie neue Computer-Tools mit heterogenen digitalen Daten die Analyse, Dokumentation und Veranschaulichung von forensischen Fällen unterstützen. Die Tools erleichtern es den Experten, auch unbekannte Zusammenhänge zu erschließen. Die Forschungsgruppe beschreibt, wie 3D-Illustrationen ebenso juristischen Laien vorgelegt werden können, um Erkenntnisse aus der Fallanalyse zu vermitteln. Wang et al. (2019) stellen in ihrer Studie ein tragbares System vor, bestehend aus einem Laserscanner, zwei Strukturlicht-Handscannern und einem kostengünstigen Virtual-Reality-Headset, um eine mehrwinklige und omnidirektionale dreidimensionale räumliche Datenerfassung von Tatorten zu ermöglichen. Das System kann genaue Informationen über Verletzungen von Verstorbenen, mögliche Verletzungswerkzeuge und Spuren am Tatort sammeln und bietet eine immersive Virtual-Reality-Darstellung der Daten. Holowko et al. (2016) schlagen eine hierarchische 3D-Scanning-Plattform vor, die eine mehrstufige Dokumentation von Tatorten ermöglicht, indem Laserscans aus verschiedenen Positionen in ein Gesamtmodell integriert werden. Der Plattform gelingt es, sowohl eine Übersicht des Tatorts als auch eine detaillierte Dokumentation von Einzelbeweisen zu liefern. Die vorgestellte Technologie ermöglicht u.a. die Virtual-Exploration von Tatorten in einer 3D-Umgebung.

3 Einordnung der Arbeit in den Fach- und Forschungsstand

Diese Arbeit stellt einen Ansatz vor, um auf Grundlage eines Videos unter Einsatz von Motion-Tracking eine real skalierte 3D-Repräsentation einer Szene zu erzeugen. Das Video soll nicht speziell für die Rekonstruktion aufgezeichnet werden müssen, stattdessen soll die Herangehensweise dazu dienen, möglichst flexibel 3D-Rekonstruktionen aus Videomaterial von einem Ereignis, wie einer Tat zu erzeugen. Dabei können Aufnahmen von Smartphones, Dashcams oder Bodycams die Datengrundlage bilden. So soll es möglich sein, die Szene als 3D-Repräsentation zu erfassen, um diese in Anwendungsgebieten der Forensik und Rechtsmedizin, beispielsweise für die Vermessung von Per-

sonen oder Objekten in Bildern oder Videos zu verwenden. Durch den im Rahmen dieser Arbeit beschriebenen Workflow soll der Prozessschritt entfallen, dass Daten direkt an einem Ereignis- oder Tatort erhoben werden müssen. Dadurch wird es möglich, 3D-Modelle von Szenen an unbekanntem Orten zu rekonstruieren und Objekte darin zu vermessen. Aufgrund des Einsatzes von Motion-Tracking soll zudem eine Alternative für die Szenenerstellung auf Basis von minderwertigem Videomaterial analysiert werden.

4 Aufbau des Workflows

Der im Rahmen einer Abschlussarbeit entwickelte Workflow zur Szenenerstellung auf Grundlage von Videomaterial und zur Größenapproximation von darin enthaltenen Objekten zeigt verschiedene Möglichkeiten auf, wie Videomaterial, durch Einsatz verschiedener Technologien (Methoden und Software) diesbezüglich, verarbeitet werden kann. Dabei haben einige Technologien spezielle Anforderungen an das vorliegende Videomaterial. Der Workflow ist hierarchisch aufgebaut und soll bei der späteren Anwendung von oben nach unten abgearbeitet werden. Hauptprozessschritte, die der Workflow beinhaltet sind: Sichten des Videomaterials, Aufbereiten des Videomaterials, 3D-Rekonstruktion einer Szene, 3D-Rekonstruktion der Form analysierender Objekte, 3D-Digitalisierung eines Objektes aus einem Video als Größenreferenz, Erfassen der Bewegung der Kamera anhand von Bildmerkmalen, Überlagerung der Größenreferenz mit dem Videomaterial, Wiederherstellung der Form der Größenreferenz zum Aufnahmezeitpunkt, Alignieren der zu analysierenden Objekte mit dem Videomaterial und Durchführung von Messungen. Jeder Hauptprozessschritt wird nochmals in Teilprozessschritten gegliedert. Beispielsweise weist der Hauptprozessschritt des *Aufbereitens des Videomaterials* als Teilprozessschritt die *Zerlegung des Videos in Einzelbilder* oder der Hauptprozessschritt der *3D-Digitalisierung eines Objektes aus einem Video als Größenreferenz* als Teilprozessschritte die Digitalisierung mittels *Structured Light Scanner* oder via *Structure from Motion* und somit durch Photogrammetrie auf. Je nach vorliegendem Videomaterial ist es möglich, entsprechende Technologien zu deren Verarbeitung zu wählen beziehungsweise Technologien auszuschließen, die aufgrund der Beschaffenheit des Videomaterials nicht verwendet werden können. Alle Technologien des Workflows verfolgen das Ziel, aus dem Videomaterial eine 3D-Szene zu erstellen, die wiederum die Grundlage für eine Objektvermessung bildet. Daher ist der letzte Prozessschritt immer die Vermessung der Objekte in den erstellten 3D-Szenen, unabhängig von der zuvor gewählten Technologie zur Verarbeitung der Daten. Bei den Messungen kann es sich um lineare Messstrecken sowie Umfangsmessungen handeln. In der vorliegenden Arbeit liegt das Hauptaugenmerk auf Prozessen zur 3D-Szenenerstellung unter Einsatz von Motion-Tracking, die besonders für Videomaterial mit minderwertiger Qualität, in Bezug z.B. auf Beleuchtung, anwendbar sind. Grund hierfür ist der Bedarf neuer Analysetechniken insbesondere für qualitativ minderwertiges Videomaterial. Aufgrund der Komplexität des Workflows werden lediglich die zuvor angesprochenen Prozesse zum Motion-Tracking im Detail

vorgestellt. In der resultierenden 3D-Szene erfolgt dann eine exemplarische Messung eines Kopfumfangs.

5 Beschreibung des Workflows auf Grundlage von Motion Tracking am Beispiel einer Kopfumfangsmessung

In diesem Kapitel wird das Vorgehen bei der Durchführung des Workflows beschrieben und somit das Zusammenspiel der einzelnen verwendeten Technologien aufgezeigt. Bei dem zugrunde liegenden Videoclip, an dem die Prozesse vorgestellt werden, handelt es sich um ein fiktiv sichergestelltes Smartphone-Video. Für die Durchführung der Messungen stand eine weibliche Versuchsperson zur Verfügung. Dieser wurde ein Maßband um den Kopf gelegt, um später an genau dieser Stelle in der Rekonstruktion den Umfang zu messen und so am digitalen Objekt erhobene Werte mit dem Realmaß vergleichen zu können.

5.1 Sichten des Videomaterials

Zu Beginn wird das Videomaterial gesichtet. Der Workflow umfasst Prozessschritte zur Rekonstruktion von 3D-Szenen aus Videos, bei denen die Kamera eine Translationsbewegung durchführt. Die Kamera muss sich von einem Ort zum anderen bewegen. Aufnahmen, die beispielsweise nur Rotationsbewegungen durchführen, bei welchen sich die Kamera auf einer Achse dreht, können nicht mit diesem Workflow bearbeitet werden. Dazu zählen insbesondere Aufnahmen, die auf einem Stativ aufgenommen wurden oder auch Aufnahmen von Überwachungskameras. Der Grund dafür ist, dass der hier vorgestellte Workflow grundsätzlich mit der Betrachtung der Parallaxe im Video (unabhängig der Videoqualität) arbeitet, welche entsteht, wenn die Kamera eine Translationsbewegung ausführt. Als Parallaxe wird die scheinbare Verschiebung eines Objekts im Verhältnis zu seiner Umgebung bezeichnet, die durch eine Änderung des Beobachtungspunkts verursacht wird. Sie wird verwendet, um die Entfernung der Kamera zu Objekten im Video zu bestimmen, indem die perspektivische Veränderung zwischen zwei verschiedenen Blickwinkeln ausgenutzt wird. Bei einer Kamerafahrt können die Frames des Videos als Einzelbilder aus mehreren Blickwinkeln interpretiert werden. Wenn keine Parallaxe entsteht, kann der Workflow nicht verwendet werden. Die zweite Einschränkung des Workflows entsteht ebenfalls durch die notwendige Parallaxe. Da die Parallaxe die scheinbare Verschiebung der Objekte zwischen den Bildern betrachtet, können Objekte, die sich während der Aufnahme in der Szene bewegen, nicht rekonstruiert werden. Bei bewegten Objekten kann die scheinbare Verschiebung nicht mehr allein auf die perspektivische Veränderung zurückgeführt werden.

5.2 Zerlegen des Videos in Einzelbilder

Der erste Schritt beinhaltet die Zerlegung des vorliegenden Videos in seine einzelnen Frames. Dazu wird die Software *Blender* verwendet. Dabei ist darauf zu achten, für die Einzelbilder ein Dateiformat zu verwenden, welches die Daten verlustfrei komprimiert, um keine Bildinformationen zu verlieren. In diesem Fall wurde das png-Format verwendet.

5.3 3D-Rekonstruktion der Szene durch händisches Motion-Tracking

In vielen Fällen kann eine 3D-Rekonstruktion einer Szene über automatische Methoden, wie Structure-from-Motion, erfolgen. Als Alternative für Anwendungsfälle, bei denen automatische Methoden nicht anwendbar sind, soll eine händische Herangehensweise beschrieben werden, um die Kamerabewegung zu rekonstruieren. In diese Kamerafahrt können anschließend durch Superimposition Objekte hineinmodelliert werden. Durch die Superimposition mit dem 3D-Scan einer Größenreferenz ergibt sich auf diese Weise auch eine real skalierte 3D-Repräsentation der Szene. Die Herangehensweise ist robuster gegenüber schlechtem Videomaterial, da hier die menschliche Auffassungsgabe genutzt wird, um Merkmale wiederzuerkennen und zu verfolgen. Mit Motion-Tracking ist es folglich möglich, Features zu extrahieren, auch wenn das Bild etwas dunkler, leicht verschwommen oder ggf. etwas schlechter aufgelöst ist. Wie bereits beschrieben, extrahiert der Structure-from-Motion-Algorithmus Bildmerkmale und versucht, diese zu einem späteren Zeitpunkt im Video wiederzuerkennen. Dabei vergleicht er Pixel miteinander, was dazu führt, dass teils schon leichte Veränderungen der Pixelwerte Probleme bei der Wiedererkennung machen. Motion-Tracking ist ein teilautomatisiertes Verfahren, das oft bei virtuellen Effekten im Bewegtbild eingesetzt wird, um 3D-Modelle in 2D-Aufnahmen einzufügen. Dabei wird ebenfalls die 2D-Position getrackt, an welcher das 3D-Modell in Superimposition überlagert werden soll. Blender stellt einige Feature für das Motion-Tracking zur Verfügung. Dazu wird zunächst die Bildsequenz in Blender importiert. Im nächsten Schritt erfolgt ein Tracking markanter Bildmerkmale über die einzelnen Frames hinweg. Dazu werden von den Nutzer:innen sogenannte Tracking-Marker in jeden Frame des Videos gesetzt. Diese Tracking-Marker müssen nun in den nachfolgenden Frames ebenfalls markiert werden. Das kann händisch geschehen, allerdings stellt Blender auch eine Reihe von automatischen Funktionen zur Feature-Suche in Frames zur Verfügung. Anschließend wird eine Vielzahl der Features getrackt. Durch das Betrachten der relativen Bewegung der frameübergreifend getrackten Features zueinander, können später durch die Parallaxe, Aussagen über deren 3D-Position getroffen werden. Abbildung 1 zeigt einen Beispielframe des Videos mit den darin getrackten Features. Beim Setzen der Marker sollten einige Kriterien beachtet werden. Zum einen sollte sich das jeweilige getrackte Objekt nicht

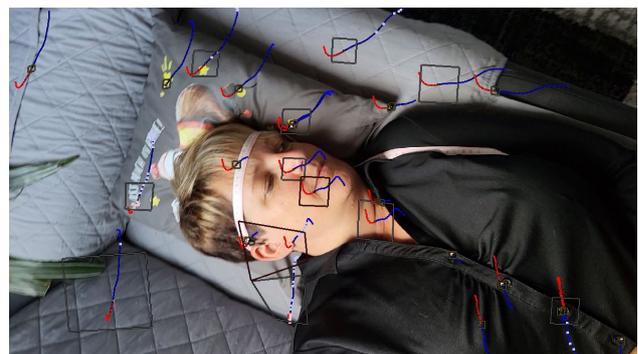


Abbildung 1: Mehrere Motion-Tracks in einem Beispielframe des Videos.

bewegen. Für die Bestimmung der 3D-Position wird die scheinbare Positionsveränderung der Features durch das Ändern der Perspektive betrachtet. Bei zusätzlicher Eigenbewegung des Objekts, wird diese Bewegung als Perspektivenwechsel fehlinterpretiert. Außerdem sollte das jeweilige Feature in vielen Frames des Videos identifizierbar sein. Für die späteren Prozessschritte ist es wichtig, dass Features extrahiert werden, die es ermöglichen, die wesentlichen Objekte in die 3D-Szene einzufügen. In diesem exemplarischen Fall zählt der Kopf, das Kissen und der Untergrund, auf dem das Kissen liegt, zu diesen Objekten von Interesse. Später findet eine 3-Punkt-Alignierung Verwendung, um die Objekte in Superimposition zum Video einzufügen. Daher werden pro Objekt mindestens drei Tracker benötigt. Die extrahierten Features können für einen 3D-Solve verwendet werden. Prinzipiell kommen hier verschiedene Algorithmen zum Einsatz, die die Bewegung der getrackten Features analysieren, um die 3D-Position und Orientierung der Kamera sowie die 3D-Position der getrackten Features in Relation zur Kamera zu berechnen. Das Ergebnis des 3D-Solves ist eine 3D-Kamera, welche die Bewegungen der tatsächlichen Kamera in der 3D-Welt nachbildet. Diese ermöglicht es, in Superimposition 3D-Elemente in das Motion-Tracking-Bild einzufügen. Abbildung 2 zeigt die rekonstruierte Kamera als Quadrat und die in Relation dazu rekonstruierten 3D-Punkte aus den getrackten Features. Dabei wurden die Features auf dem Kissen grün eingefärbt, die Features auf dem Kopf rot, und die Features auf dem Untergrund orange.

5.4 Händische 3D-Rekonstruktion von Objekten in der Kamerafahrt

Die rekonstruierte Punktvolke und die Kamerafahrt können nun zum Einfügen der relevanten Objekte in die Szene verwendet werden. Um den Untergrund, auf dem das Kissen liegt, zu rekonstruieren, werden die 3D-Punkte der jeweiligen getrackten Features auf dem Untergrund vermascht, sodass eine Fläche entsteht.

5.5 Geeignetes Objekt für die Größenreferenz im Video finden

Um die 3D-Szene später mit dem Realmaßstab anzureichern, muss ein Objekt im Video gefunden werden, welches als Größenreferenz dienen kann. Dabei sollte das Objekt folgende Bedingungen erfüllen:

- Das Objekt sollte physisch zugänglich sein, da in späteren Prozessschritten ein 3D-Scan dieses Objektes erzeugt werden muss.
- Das Objekt darf sich im Video nicht bewegen, da aufgrund der Parallaxe nur unbewegte Objekte rekonstruiert werden können.
- Das Objekt sollte im Video gut zu erkennen sein. Es sollte möglichst scharf und möglichst über einen großen Zeitraum im Video zu sehen sein.
- Die Materialeigenschaften des Objektes sollten sich für einen 3D-Scan eignen. Die genauen Anforderungen hier abhängig von der jeweiligen verwendeten Scan-Methode.

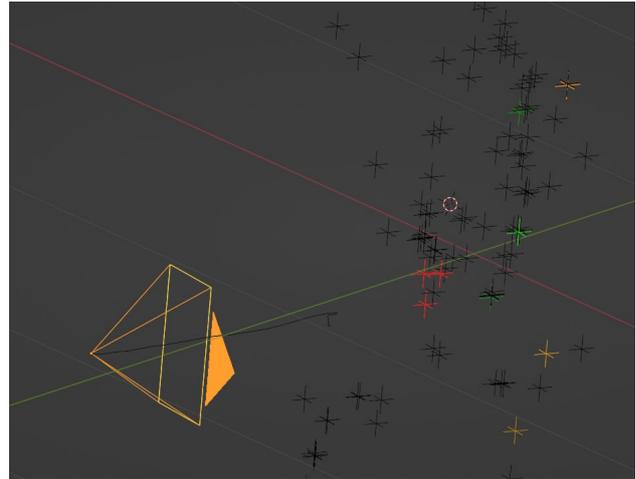


Abbildung 2: Rekonstruierte Kamerafahrt und Motion-Tracks im 3D-Raum.

Bei der Auswahl der Größenreferenz muss zwischen der Sichtbarkeit im Video und der Eignung für einen 3D-Scan abgewogen werden. Allerdings können auch mehrere Objekte als Größenreferenz gewählt werden. Die Eignung für den Scan ist abhängig von dem Material beziehungsweise Oberflächeneigenschaften des Objektes. In diesem Workflow wird als Scan-Methode Structure-from-Motion (Iglhaut, 2019) vorgestellt. Dabei handelt es sich um eine optische Scan-Methode aus dem Bereich der Photogrammetrie. Optische Scan-Methoden weisen in der Regel Probleme mit stark reflektierenden, transparenten oder homogenen Oberflächen auf. Als Größenreferenz dient hier



Abbildung 3: A: Darstellung des Kissens aus dem Video als exemplarische Größenreferenz B: Versuchsperson mit dem Maßband um die Stirn.

das Kissen, auf welchem die Versuchsperson im Video liegt. Abbildung 3 zeigt das verwendete Kissen und die Versuchsperson.

5.6 3D-Scan des Größenreferenzobjektes mit Spiegelreflexkameras und Structure-from-Motion

Sollte kein 3D-Scanner für den Scan zur Verfügung stehen, kann das Objekt, welches als Größenreferenz dient, mit handelsüblichen Spiegelreflexkameras digitalisiert werden. Dabei wird zunächst eine 360°-Aufnahme von dem zu scannenden Objekt angefertigt. Hierbei dient ein Glastisch als Unterlage und Halterung des Kissens. Der Glastisch steht dabei auf einem Drehteller, welcher mit einem PC angesteuert wird. Es werden drei Kameras verwendet, um das Objekt aus drei unterschiedlichen Höhen aufzunehmen. So entstehen eine Draufsicht, eine frontale Ansicht und eine Ansicht von unten. Dabei ist zu beachten, dass die Kameras aus den verschiedenen Blickwinkeln zu etwa 80 % den gleichen Bildbereich abdecken, da das 3D-Modell durch Structure-from-Motion aus den Bildern erzeugt werden soll. Daher muss der jeweilige Bildbereich in den anderen Bildbereichen zu sehen sein, um die Verbindung herzustellen. Aus demselben Grund wird bei der 360°-Aufnahme bei zwei Grad Drehung des Drehtellers ein Bild mit allen drei Kameras aufgenommen. Das Objekt wird bei der Aufnahme von mehreren Softboxen ausgeleuchtet. Abbildung 4 zeigt das Scan-Setup. Die 360°-Aufnahme simuliert eine Kamerafahrt um das Objekt. Während der Berechnung werden die Blickwinkel der einzelnen Bilder auf das Objekt in Form von Positionskoordinaten und Rotationskoordinaten berechnet. Daher sollte darauf geachtet werden, dass die meisten Texturinformationen im Bild auf dem zu rekonstruierenden Objekt liegen und nicht im



Abbildung 4: Scan-Setup für die Aufnahme der Bilder zur 3D-Digitalisierung des Kissens mithilfe von Structure-from-Motion.

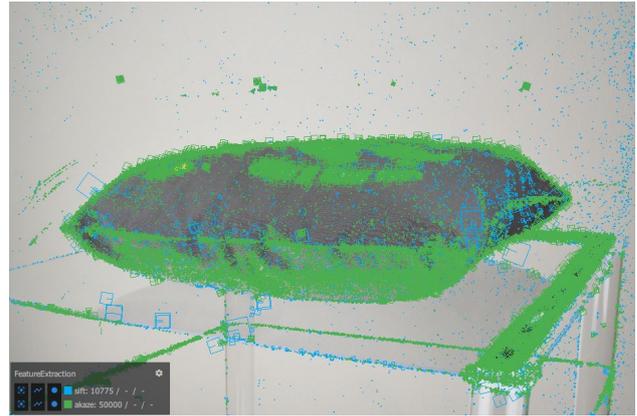


Abbildung 5: Mit dem Template extrahierte SIFT und AKAZE Features am Beispiel eines Bildes aus der Bilderserie der 3D-Digitalisierung des Kissens.

Hintergrund oder im Vordergrund. Die Bilder sollten also vor einem möglichst homogenen Hintergrund aufgenommen werden. Im Ergebnis werden verschiedene Feature, wie SIFT und AKAZE (Yamada, 2018), berechnet, die die Grundlage für die Erstellung einer 3D-Oberfläche bilden. In der Nachfolgenden Abbildung 5 sind das Kissen und berechnete Feature dargestellt. Bei den Aufzeichnungsparametern ist darauf zu achten, dass eine möglichst niedrige Sensorverstärkung (ISO-Wert) und eine möglichst weit geschlossene Blende gewählt werden. Das Ziel ist es, komplett scharfe Bilder ohne Rauschen zu erzeugen. Die Sensorverstärkung erzeugt Rauschen in den Farbwerten der Pixel, was zu vermeiden ist, und eine möglichst kleine Blendenöffnung vermeidet Tiefenunschärfe. Diese beiden Einstellungen sorgen allerdings dafür, dass wenig Licht auf den Sensor trifft und die Aufnahmen sehr dunkel sind. Um dem entgegenzuwirken, wird das Objekt stark ausgeleuchtet und es wird eine hohe Belichtungszeit gewählt. Eine hohe Belichtungszeit erzeugt allerdings Bewegungsunschärfe bei bewegten Objekten. Daher wird der Drehteller für das Aufnehmen der einzelnen Bilder angehalten. Neben dem eigentlichen Objekt befinden sich ebenfalls zwei Concentric Circles Tags (CCTags) auf den Bildern. Der Abstand zwischen den beiden Tags ist bekannt, wird in der Photogrammetriesoftware Meshroom hinterlegt und bei der Rekonstruktion als Größenreferenz verwendet. Bei der Berechnung des 3D-Modells des Kissens mittels Structure-from-Motion wird ein CCTag verwendet, um das Modell in den realen Maßstab zu bringen. Abbildung 6 zeigt das texturierte polygonale 3D-Modell des Kissens.

5.7 Alignieren der beliebig skalierten 3D-Szene mit der Größenreferenz

Nun wird die Szene mit dem Scan der Größenreferenz angereichert. In der durch Motion-Tracking entstandenen 3D-Punktwolke werden zunächst die drei getrackten Punkte auf dem Kissen zu einem Dreieck vermascht. Jetzt kann die Alignierung des Dreiecks durch die 3-Punkt-Alignierung mit dem real skalierten Scan der Größenreferenz erfolgen. Hierfür werden markante Punkte benötigt. Bei den markanten Punkten handelt es sich um jene, die durch Motion-Tracking aus dem Videomaterial extrahiert wurden und jetzt das vermaschte Dreieck aufspannen.



Abbildung 6: Durch Structure-from-Motion berechnetes realskaliertes 3D-Modell des Kissens.

Diese Punkte werden anschließend mit den korrespondierenden Punkten auf dem real skalierten Scan verbunden. Somit ergibt sich die Superimposition der real skalierten Größenreferenz mit der Kamerafahrt.

5.8 Dummyobjekte in die Szene einfügen

In die nun real skalierte 3D-Szene können weitere Objekte in Superimposition zum Video hineinmodelliert werden. In der Motion-Tracking rekonstruierten Szene ist dies notwendig, um einen Dummykopf zu erzeugen, der es erlaubt, den Kopfumfang zu messen. Hierzu wird erneut ein Dreieck erzeugt, indem die Punkte des Kopfes aus dem Motion-Tracking vermascht werden. Dieses Dreieck wird verwendet, um durch eine 3-Punkt-Alignierung ein Dummymodell eines Kopfes in die Szene einzufügen. Hier wurde ein nachmodellierter menschlicher Kopf verwendet. Abbildung 7 zeigt das Resultat der Alignierung. In Abbildung 8 wird das Resultat dargestellt, wie das Kissen, der Kopf und der Untergrund in der Punktwolke liegen. Zur besseren Darstellung wurde hier das Video ausgeblendet. Die Superimposition wird jetzt verwendet, um Feineinstellungen vorzunehmen. Beispielsweise kann der Dummykopf mit den Modellierungswerkzeugen in Blender in seinen Proportionen so angepasst werden, dass er mit jedem 2D-Abbild des Kopfes in jedem Frame in Superimposition übereinstimmt.



Abbildung 7: Resultat der Alignierung des Dummykopfes mit der real skalierten 3D-Szene aus dem Motion-Tracking.

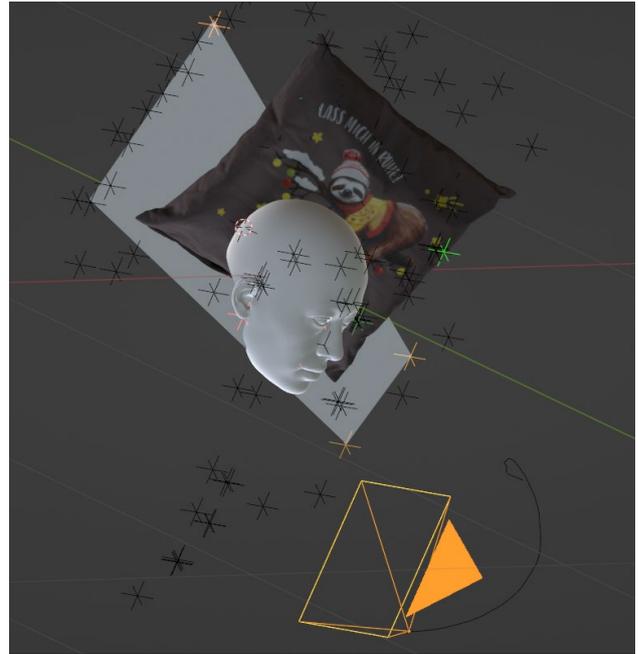


Abbildung 8: Kissen, Kopf und Untergrund in der rekonstruierten Kamerafahrt und Punktwolke der Motion-Tracks.

5.9 Messungen in der rekonstruierten 3D-Szene

In dieser rekonstruierten 3D-Szene können nun Messungen erfolgen. Im Rahmen dieser Arbeit wird eine Messmethode vorgestellt, um den Umfang des Kopfes zu approximieren, welche als Bézier bezeichnet wird. Bei dieser Messmethode wird eine Bézierkurve in Superimposition an die Stelle des Videos eingefügt, an der der Umfang des Objektes gemessen werden soll. Dazu erfolgt zunächst eine Verschiebung des zu messenden Kopfes in den Ursprung des globalen Koordinatensystems von Blender. Da die Rückseite des Kopfes auf dem Kissen liegt und nicht im Videomaterial einsehbar ist, wird die Annahme getroffen, dass die linke und rechte Kopfseite symmetrisch sind. Nun wird der erste Teil der Bézierkurve erzeugt und auf die zweite Gesichtshälfte gespiegelt. So entsteht eine vollständige Instanz der Bézierkurve, die zur Vermessung des Kopfumfanges verwendet wird. Die Superimposition dient zur Überprüfung und Anpassung der Bézierkurve auf dem zu

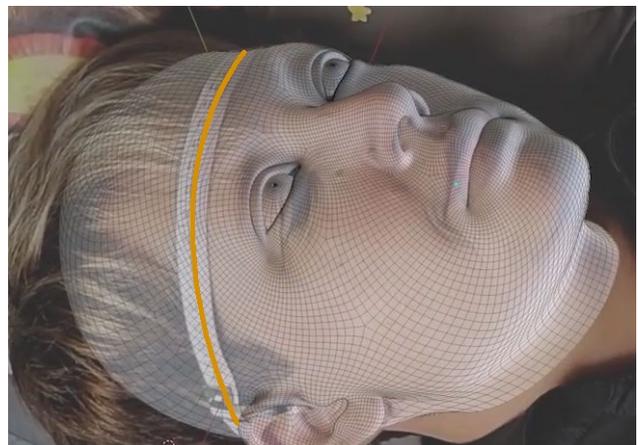


Abbildung 9: Superimposition der Bezierkurve an dem Dummykopf mit Frame des Videos.

vermessenden Objekt. Dazu wird jeder Frame der Bildsequenz betrachtet. Abbildung 9 zeigt einen beispielhaften Frame dieser Superimposition.

Anschließend wird die Länge der Bézierkurve gemessen. In dieser exemplarischen Messung wurde ein Umfang von 60,8 cm gemessen. Das Maßband im Video misst einen physischen Umfang von 58,2 cm.

6 Zusammenfassung

Der vorgestellte Workflow bietet verschiedene Technologien, Verfahren und Techniken zur 3D-Rekonstruktion von Szenen. Das Hauptziel besteht darin, Objekte in diesen Szenen zu vermessen, um forensische, kriminalistische und rechtsmedizinische Fragestellungen zu beantworten. Der Workflow umfasst mehrere Schritte, von der Erfassung der Szene bis zur Analyse der rekonstruierten 3D-Daten. Jeder Schritt des Workflows birgt potenzielle Fehlerquellen, die genauer analysiert und optimiert werden müssen. Trotzdem zeigt sich bereits jetzt anhand des hier vorgestellten Beispiels, dass der Gesamtprozess als unterstützende Methode für herkömmliche Analysen eingesetzt werden kann. Vor allem Motion-Tracking stellt eine vielversprechende Alternative zu bekannten Verfahren wie Structure-from-Motion dar. Je nach Fragestellungen, Ressourcen und Zeit können entsprechende Verarbeitungsschritte für vorliegendes Videomaterial gewählt werden. Abhängig davon ergeben sich jedoch die bereits angesprochenen Fehlerquellen, wie beispielsweise Verzerrungen oder Ungenauigkeiten in den Aufnahmen, die die Qualität der rekonstruierten Szene beeinflussen. Auch bei der Vermessung der Objekte können Messfehler auftreten, die zu Ungenauigkeiten führen können. Um die Qualität des Workflows zu verbessern, sollten diese Fehlerquellen analysiert und gegebenenfalls optimiert werden. Dies könnte beispielsweise durch die Kalibrierung der Kameras oder die Verwendung hochpräziser Messinstrumente geschehen. Darüber hinaus könnte die Anwendung von Algorithmen zur automatischen Fehlererkennung und -korrektur helfen, Ungenauigkeiten zu reduzieren. Eine ausführliche SWOT-Betrachtung des Workflows könnte ebenfalls einen Mehrwert liefern.

7 Bedeutung und Potential des Workflows für die Ermittlungsarbeit

Trotz der potenziellen Fehlerquellen, die im vorherigen Kapitel angesprochen wurden, zeigt sich bereits jetzt, dass der vorgestellte Workflow eine wertvolle Ergänzung zu herkömmlichen Analysemethoden darstellen kann. Die 3D-Rekonstruktion von Szenen mittels Motion-Tracking ermöglicht eine detaillierte Vermessung von Objekten, was forensischen, kriminalistischen und rechtsmedizinischen Fachleuten dabei hilft, genaue Informationen zu gewinnen. Vor allem im Alltag agierender Strafverfolgungsbehörden weist zu Grunde liegendes Analysematerial zumeist eine eher minderwertige Bild- und Videoqualität auf. Hier gilt es dann insbesondere, daraus so viel Informationen wie nur möglich für weiterführende Analysen zu gewinnen. Dabei können Workflows, wie der hier vorgestellte, unterstützend eingesetzt werden. Ergänzend dazu sei zu erwähnen,

dass selbst mit bereits vorhandenem Equipment aus dem Ermittlungsalltag, wie Spiegelreflexkameras und Desktop-PCs, Technologien aus dem Workflow angewandt werden können. Hochpreisige Ausstattung und Experten-Know-How sind nicht zwangsläufig Voraussetzung für die erfolgreiche Anwendung des Workflows.

Literaturverzeichnis

- Bornik, A., Urschler, M., Schmalstieg, D., Bischof, H., Krauskopf, A., Schwark, T., ... & Yen, K. (2018). Integrated computer-aided forensic case analysis, presentation, and documentation based on multi-modal 3D data. *Forensic science international*, 287, 12-24. DOI: 10.1016/j.forsciint.2018.03.031
- Buck, U., Naether, S., Braun, M., Bolliger, S., Friederich, H., Jackowski, C., ... & Thali, M. J. (2007). Application of 3D documentation and geometric reconstruction methods in traffic accident analysis: with high resolution surface scanning, radiological MSCT/MRI scanning and real data based animation. *Forensic science international*, 170(1), 20-28. DOI: 10.1016/j.forsciint.2006.08.024
- Campana, L., Breitbeck, R., Bauer-Kreuz, R., & Buck, U. (2016). 3D documentation and visualization of external injury findings by integration of simple photography in CT/MRI data sets (IprojeCT). *International journal of legal medicine*, 130, 787-797. DOI: 10.1007/s00414-015-1274-3
- Hołowko, E., Januszkiewicz, K., Bolewicki, P., Sitnik, R., & Michoński, J. (2016). Application of multi-resolution 3D techniques in crime scene documentation with bloodstain pattern analysis. *Forensic science international*, 267, 218-227. DOI: 10.1016/j.forsciint.2016.08.036
- Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., & Rosette, J. (2019). Structure from motion photogrammetry in forestry: A review. *Current Forestry Reports*, 5, 155-168. DOI: <https://doi.org/10.1007/s40725-019-00094-3>
- Wang, J., Li, Z., Hu, W., Shao, Y., Wang, L., Wu, R., ... & Chen, Y. (2019). Virtual reality and integrated crime scene scanning for immersive and heterogeneous crime scene reconstruction. *Forensic science international*, 303, 109943. DOI: 10.1016/j.forsciint.2019.109943
- Yamada, K., & Kimura, A. (2018, January). A performance evaluation of keypoints detection methods SIFT and AKAZE for 3D reconstruction. In 2018 International Workshop on Advanced Image Technology (IWAIT) (S. 1-4). IEEE. DOI: 10.1109/IWAIT.2018.8369647