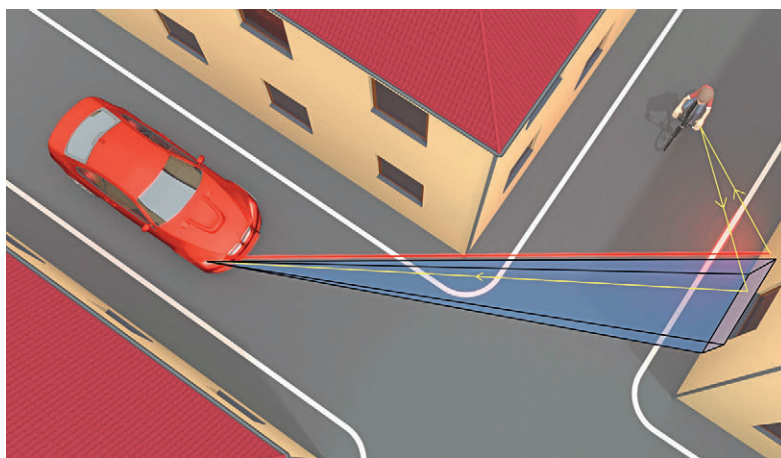


# Wenn eine Wand kein Hindernis mehr ist

Ein noch sehr junges Teilgebiet der Computer-Vision beschäftigt sich mit der Lokalisierung und Verfolgung von Objekten, die sich nicht im direkten Sichtfeld einer Kamera befinden. Bisher waren dazu teure Spezialkameras und Laserbeleuchter mit einer hohen zeitlichen Auflösung erforderlich. Ein neuer Ansatz löst dieses Problem nun mit handelsüblicher Hardware.



**Bild 1:** Potenzielle Verbesserung für die Verkehrssicherheit: Eine drohende Kollision mit einem Verkehrsteilnehmer außerhalb des Sichtbereichs kann rechtzeitig erkannt werden.

Jonathan Klein<sup>1,2</sup>, Martin Laurenzis<sup>1</sup>,  
Matthias B. Hullin<sup>2</sup>  
<sup>1</sup>Deutsch-Französisches Forschungs-  
institut Saint-Louis  
<sup>2</sup>Universität Bonn

Optische Systeme sind das wohl wichtigste Werkzeug zur Wahrnehmung der unmittelbaren Umgebung. Eine direkte Sichtverbindung zu einem Objekt ermöglicht dabei das einfache und schnelle Sammeln einer Vielzahl relevanter Informationen wie die Position, Größe, Typ oder Bewegungsrichtung des Objekts. Auch der Mensch vertraut in vielen alltäglichen Aufgaben hauptsächlich auf seine Augen als Informationsquelle.

Allerdings kann es leicht passieren, dass die Sichtverbindung unterbrochen und eine visuelle Wahrnehmung dadurch unmöglich wird. Ein System, das diese Limitierungen ganz oder zumindest teilweise umgehen kann, würde in solchen Situationen einen bedeutenden Vorteil darstellen. Solche Systeme sind

Gegenstand aktueller Forschung in einem Teilgebiet der Computer-Vision: Non-Line-of-Sight (NLoS)-Imaging oder die optische Erfassung von Objekten außerhalb des Sichtfeldes.

Anwendungsgebiete für solche Systeme finden sich zum Beispiel in autonomen oder teilautonomen Fahrzeugen. Drohende Kollisionen auch mit nicht direkt sichtbaren Fahrzeugen vermeiden zu können, würde im zunehmenden Straßenverkehr einen entscheidenden Beitrag zur Verkehrssicherheit leisten (**Bild 1**). Auch in Rettungseinsätzen, wie z. B. bei Bränden, ermöglicht indirektes Sehen die Lokalisation von gefährdeten Menschen, ohne dass sich die Rettungskräfte derselben Gefahr aussetzen müssen (**Bild 2**). Generell ist ein solches System überall dort interessant, wo eine direkte Sichtverbindung entweder nicht erreichbar oder zu gefährlich wäre.

NLoS-Imaging ist ein noch recht junges Forschungsgebiet, das erst 2009 von Wissenschaftlern an den US-amerikanischen MIT Media Labs begründet

wurde [1]. In den letzten Jahren gab es zahlreiche weitere Veröffentlichungen auf diesem Gebiet, die die Rekonstruktion teilweise beachtlich verbessert haben [2]. Insgesamt handelt es sich jedoch um Grundlagenforschung und kommerzielle Produkte sind noch nicht verfügbar.

## Diffuse Spiegel

Das Problem der fehlenden Sichtverbindung lässt sich mit Spiegel auf triviale Art lösen, beispielsweise durch die Benutzung von Periskopen. In vielen Fällen sind Spiegel allerdings nicht verfügbar, weswegen neue Ansätze gesucht werden.

Betrachtet man das zugrunde liegende Prinzip, lösen Spiegel das Problem der fehlenden Sichtverbindung, indem sie durch Reflexionen die Sichtstrahlen umlenken. Allerdings reflektieren nicht nur Spiegel das Licht, sondern auch alle anderen sichtbaren Materialien. Der Unterschied besteht darin, dass im Gegensatz zu Spiegeln bei diffusen Materialien das Licht nicht in eine einzige Richtung reflektiert, sondern auch mit Licht aus allen anderen Richtungen gemischt wird und somit sämtliche Richtungsinformationen verloren gehen. Die Herausforderung des NLoS-Imaging besteht nun darin, aus diesen diffusen Reflexionen dennoch sinnvolle Informationen zu errechnen.

In **Bild 3** wird das grundlegende Prinzip dargestellt. In diesem Szenario befindet sich das Kamerasystem auf der rechten Seite, während das Zielobjekt auf der linken Seite ist und von einer Wand verdeckt wird. Um eine indirekte Sichtverbindung aufzubauen, wird ein Laser auf die Rückwand gerichtet, der von dieser in alle Richtungen reflektiert wird. Ein Teil des Lichts trifft auf das Objekt im Raum, wo eine zweite Reflexion stattfindet.

**Bild 2:** Ein weiteres mögliches Einsatzgebiet für NLoS-Imaging: Eine Person wird in einem brennenden Haus lokalisiert.



det. Eine dritte Reflexion an der Rückwand ermöglicht es schließlich, dass das Licht seinen Weg zurück in die Kamera findet und ein Signal gemessen werden kann. Die Rückwand fungiert also als diffuser Spiegel.

Die Rückwand in dieser Abbildung zeigt die Intensitätsverteilung des von dem Objekt zur Wand zurückgeworfenen Lichts. Aufgrund der diffusen Natur der Reflexionen ist das Signal stark verschwommen und eine direkte Rekonstruktion ist nicht möglich. Die transiente Lichtmessung bietet eine Möglichkeit, um die Bildinformationen zu extrahieren.

**Transiente Lichtmessung**

Wenn eine Szene beleuchtet wird, reflektiert das Licht zwischen den verschiedenen Objekten hin und her, bis sich ein Gleichgewicht einstellt. Aufgrund der Lichtgeschwindigkeit spielt sich dieser Effekt im Bereich von Nanosekunden ab und ist mit herkömmlicher Technik nicht messbar. Spezielle Kameras können diesen flüchtigen (Engl.: transient) Prozess jedoch mithilfe einer mit der Kamera synchronisierten Lichtquelle aufzeichnen.

Durch den zeitlichen Unterschied zwischen Erzeugung und Messung von Lichtpulsen kann dann für jeden Kamerapixel eine Entfernung gemessen werden, indem über die bereits bekannte Lichtgeschwindigkeit der Zeitunterschied in die zurückgelegte Wegstrecke umgerechnet wird. Diese Entfernungsmessung wird unter anderem von Spielkonsolen zur Gestenerkennung verwendet.

**Rekonstruktion verdeckter Szenen**

Im Falle von NLoS-Imaging ist es aber nicht das primäre Ziel, die Entfernung zu dem diffusen Reflektor zu bestimmen. Die einzelnen Kamerapixel enthalten in diesem Fall nicht nur Entfernungsinformationen über den Reflektor, sondern auch über die verdeckten Objekte. Durch diese zeitliche Auflösung der Daten kann die verlorengangene Richtungsauflösung durch die diffusen Reflexionen nun teilweise kompensiert werden.

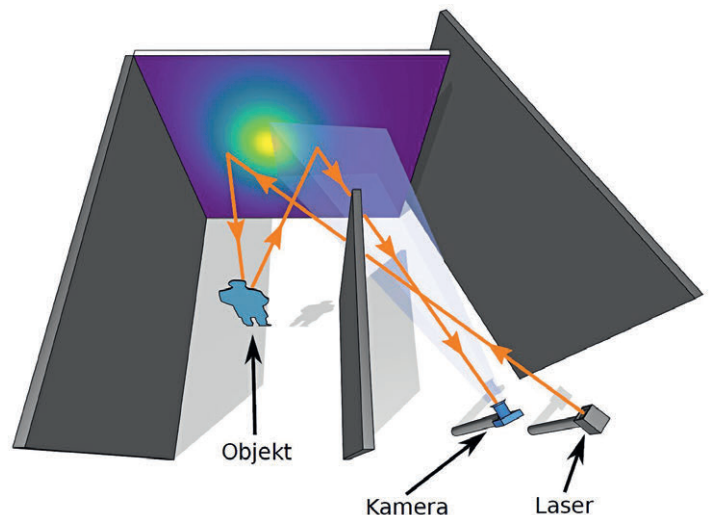
Wie in **Bild 4** gezeigt, kann zur Rekonstruktion verdeckter Objekte das aus der Tomographie bekannte Verfahren der Rückprojektion verwendet werden. Der Laserstrahl trifft die Wand an Punkt L, der durch eine vorherige Kalibration bekannt ist. Im hellblauen Sichtfeld der Kamera wird das von dem Objekt reflektierte Licht an verschiedenen Punkten (hier A und B) gemessen. Durch die transiente Messung ist die Zeitdifferenz

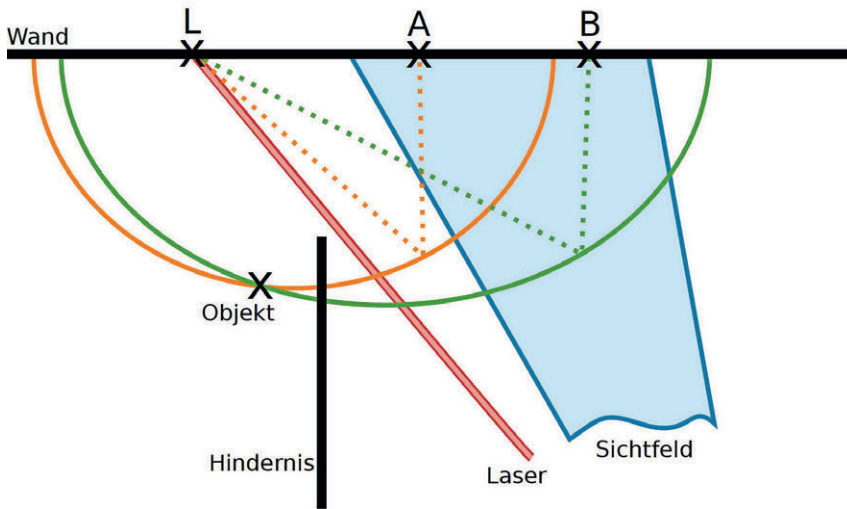
und damit die Strecke von Punkt L zu Punkt A und B (gestrichelte Linien) bekannt, jedoch nicht die Objektposition, an der das Licht reflektiert wurde. Alle möglichen Objektpositionen die zu den Laufzeitmessungen passen, liegen jedoch auf einer Ellipse (durchgezogene Linien), die durch die Ellipsenbrennpunkte L und A bzw. B definiert ist. Für jeden Messpunkt auf der Wand und damit für jedes Sensorelement der Kamera erhält man nun eine leicht unterschiedliche Ellipse. Da sich das Objekt an einer Position auf allen diesen Ellipsen befinden muss, kommt als Objektposition nur der gemeinsame Schnittpunkt aller Ellipsen in Frage. Die Rückprojektion stellt somit eine effektive Möglichkeit zur Rekonstruktion dar, setzt aber zwingend eine hohe zeitliche Auflösung voraus. In einer Kooperation zwischen dem Deutsch-Französischen Forschungsinstitut ISL und der Universität Bonn wurde nun ein neuartiges Verfahren entwickelt das mit nahezu beliebigen Kameradaten arbeiten kann [3].

**„Analyse-durch-Synthese“**

Die Kernidee des Verfahrens basiert auf einer alten Beobachtung der Mathematik. Es gibt zahlreiche Probleme, die einfach zu lösen, aber sehr schwierig umzukehren sind. So ist es beispielsweise einfach, zwei große Primzahlen zu multiplizieren, aber es ist nicht effizient möglich, das Produkt wieder in seine Primfaktoren zu zerlegen. Auf dieselbe Weise stellt sich heraus, dass zu NLoS-Imaging ein Umkehrproblem existiert, das gut erforscht und verhältnismäßig effizient zu lösen ist: Das Signal aus diffusen Reflexionen, das ein Objekt ➤

**Bild 3:** Der grundlegender Aufbau: Durch mehrere diffuse Reflexionen wird eine Sichtverbindung zwischen dem Objekt und der Kamera geschaffen.



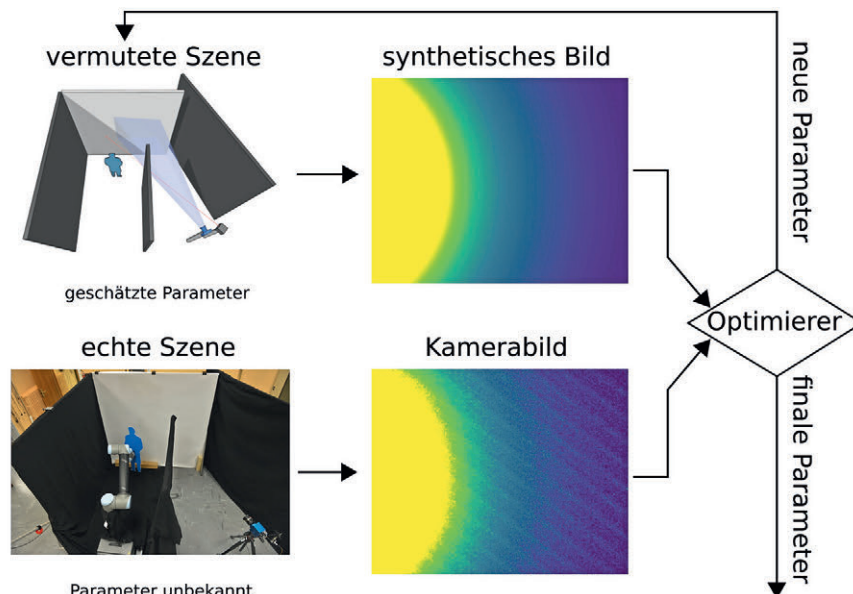


**Bild 4:** Rekonstruktion durch Rückprojektion: Die Distanzmessungen der Messpunkte A und B definieren jeweils eine Ellipse auf deren Schnittpunkt sich das Objekt befinden muss.

auf den Reflektor zurückwirft, lässt sich mit gängigen Techniken der Computer-Graphik akkurat simulieren. Diese „Vorwärts-Simulation“ von Lichtausbreitung ist seit vielen Jahren Gegenstand aktiver Forschung und ist ein Standardwerkzeug in der Film- und Computerspielproduktion.

**Bild 5** zeigt, wie man durch Zuhilfenahme von numerischer Optimierung das ursprüngliche Problem der Rückwärts-Simulation („woher kommt das Licht?“) durch sein inverses Äquivalent, die Vorwärts-Simulation, lösen kann.

Möchte man die Position eines Objekts bestimmen, wird diese zunächst grob geschätzt, beispielsweise durch den Szenenmittelpunkt. Zusammen mit einer vorherigen Messung der relativen Position und Ausrichtung der diffus reflektierenden Wand (die Szenenkalibration) kann dann das Bild berechnet werden, das die Kamera auf der Wand „sieht“, wenn sich das Objekt tatsächlich an der geschätzten Position befinden würde. Zur gleichen Zeit wird ein Bild der echten Szene aufgenommen, das nun mit dem synthetischen Bild verglichen wer-



**Bild 5:** Das Analyse-durch-Synthese-Prinzip: Das Kamerabild wird mit dem berechneten Bild verglichen und durch numerische Optimierung der Szenenparameter angepasst.

den kann. Die geschätzte Objektposition wird nun iterativ angepasst, bis das synthetische Bild dem echten Kamerabild entspricht und die Objektposition gefunden wurde.

Für den Vergleich werden beide Bilder als Vektoren interpretiert und die Minimierung ihrer Vektordifferenz als Least-Squares-Optimierungsproblem formuliert. Diese Art von Problem kann dann mit etablierten numerischen Verfahren gelöst werden. Dieser Ansatz ist extrem vielseitig und leicht erweiterbar, da das eigentlich komplexe NLoS-Imaging-Problem auf ein Bildvergleichsproblem reduziert wird. Da allgemeine Renderverfahren nahezu beliebige Szenen künstlich erzeugen können, bedeutet dies, dass sich prinzipiell auch beliebige Szenenparameter (zum Beispiel die Position, Größe, Ausrichtung, Form, Textur und das Material von Objekten) durch diese Technik rekonstruieren lassen. Um das Rendering zu ermöglichen, müssen nicht rekonstruierte Szenenparameter vorab allerdings entweder bekannt sein oder sinnvoll geschätzt werden.

In der Praxis stößt man dabei jedoch auf eine Einschränkung: Wächst die Zahl der zu rekonstruierenden Szenenparameter, erhöht sich auch die Anzahl der Freiheitsgrade in der numerischen Optimierung, was diese langsamer und fehleranfälliger macht. Wenn Szenenparameter das gemessene Signal nur marginal beeinflussen (wie es etwa kleine Details in der Objektgeometrie tun), steigen damit auch die Anforderungen an die Kalibration des Messsystems und die Wahrscheinlichkeit wächst, dass nicht die optimale Lösung gefunden wird. Auch die Rechenzeit kann rapide anwachsen, da durch die zusätzlichen Parameter nicht nur jeder einzelne Iterationsschritt aufwendiger zu berechnen wird, sondern durch die gestiegene Komplexität auch insgesamt mehr Iterationsschritte benötigt werden.

In vielen Fällen ist eine vollumfängliche Rekonstruktion allerdings überhaupt nicht erforderlich. In dem zu Beginn erwähnten Szenario des Hausbrands ist es beispielsweise wesentlich interessanter zu wissen, in welchem Raum sich wie viele Personen aufhalten, als die Farbe ihrer Schuhe oder die Länge ihrer Haare zu rekonstruieren. In Szenarien wie diesen macht sich die Skalierbarkeit des neuen Ansatzes bezahlt: Eine einfache Positionsbestim-





**Bild 6:** Die verwendete Hardware: Xenics Xeva-1.7-320 Infrarot-Kamera und SemiNex 4PN-108 1550 nm Laserdiode

mung kann mit hoher Robustheit in Echtzeit durchgeführt werden.

Dieses Verfahren ist prinzipiell unabhängig von der verwendeten Hardware, solange es eine (durch Kalibrierung gewonnene) Funktion gibt, die aus den berechneten Lichtintensitäten synthetische Kamerabilder erzeugt. Dabei können grundsätzlich auch Tiefenbilder oder die von AMCW-LIDAR Kameras generierten Phasenbilder erzeugt und mit der echten Aufnahme verglichen werden. Tatsächlich stellt sich jedoch heraus, dass das Verfahren nicht zwingend auf zeitliche Informationen angewiesen ist, sondern sie eine rein optionale zusätzliche Informationsquelle darstellen. In der Praxis kann eine höhere Rauschanfälligkeit dieser Kameras tatsächlich von größerer Bedeutung sein, als die zusätzliche Information aus den Tiefendaten, sodass das System mit handelsüblichen Intensitätskameras bessere Ergebnisse liefert.

### Erste Experimente sind vielversprechend

Bisher steckt die neue Technik noch in den Kinderschuhen, und es sind noch keine kommerziellen Produkte erhältlich. Die erste Hürde der experimentellen Verifikation wurde aber bereits gemeistert: In einem Experiment wurden Infrarot-Kamera und Laser (**Bild 6**) in einem Raum von etwa 2 x 2 m<sup>2</sup> Größe platziert und ein Objekt mit einer Frequenz von ca. 10 Hz und einer Genauigkeit von unter 5 cm verfolgt. Eine Wel-

lenlänge von 1550 nm wurde gewählt, da sie nicht nur unsichtbar, sondern auch in höheren Intensitäten augensicher ist. Dadurch kann das Rauschverhalten des gemessenen Signals deutlich verbessert werden.

Die Experimente zeigen, dass sich durch eine sorgfältige Kalibration die Ergebnisse maßgeblich verbessern lassen. In Zukunft wäre daher ein System, dass sich vollautomatisiert auf neue Begebenheiten kalibrieren kann, wünschenswert. Zusätzlich zeigt sich aber auch, dass fehlende Szeneninformationen in vielen Fällen durch konservative Schätzungen gut angenähert werden können. Das System kann daher mit zusätzlichen Informationen (zum Beispiel die Geometrie des Raumes, in dem sich das Objekt befindet) zwar genauere Ergebnisse liefern, bleibt ohne diese Informationen aber weiterhin funktionsfähig (in diesem Beispiel verschlechterte sich die Genauigkeit um einen systematischen Fehler von ca. 10 - 15 cm).

In weiteren Experimenten muss sich nun zeigen, wo die Grenzen des Verfahrens liegen. Es wird erwartet, dass komplexere Szenen mit mehreren Objekten oder Objekten, die sich sehr weit von der Wand weg befinden, gewisse Schwierigkeiten bereiten und eine Rekonstruktion vielleicht sogar unmöglich machen könnten.

Auch stellt sich die Frage, inwieweit sich nicht nur die Position, sondern auch die Geometrie von Objekten rekonstruieren lässt, oder ob man das mathematische Optimierungsproblem durch eine andere Formulierung robuster und schneller lösen könnte. Eine genauere Untersuchung der Unterschiede zwischen Analyse-durch-Synthese und Rückprojektion mit besonderem Hinblick auf Hardwarevoraussetzungen und erwarteter Genauigkeit steht zurzeit ebenfalls noch aus.

### Fazit

NLoS-Imaging ist ein junges und aufstrebendes Forschungsgebiet, in dem bisher ungenutzte Information diffuser Reflexionen auf eine neue Art analysiert werden und damit eine Rekonstruktion verdeckter Objekte ermöglicht. Vergleichbar mit Big-Data-Algorithmen, die in scheinbar strukturlosen Daten Zusammenhänge erkennen können, erschließt sich mit NLoS-Imaging eine vollkommen neue Art der optischen Messung, deren volles Potenzial erst Stückweise erkannt wird und die Messtechnik nachhaltig verändern könnte.

Die Analyse durch Synthese stellt einen orthogonalen Ansatz zur bisher üblichen Rückprojektion zur Lösung von NLoS-Problemen dar. Die niedrigeren Hardwareanforderungen machen diese Technik für eine Vielzahl von Szenarien interessant und die Experimente zeigen, dass die Leistung mit der bisher üblichen Rückprojektion problemlos mithalten kann. ■

### Kontakt

Jonathan Klein  
ISL  
Universität Bonn  
kleinj@cs.uni-bonn.de



Dr. Martin Laurenzis  
ISL  
Martin.Laurenzis@isl.eu



Prof. Dr. Matthias B. Hullin  
Universität Bonn  
hullin@cs.uni-bonn.de



Deutsch-Französisches  
Forschungsinstitut Saint-Louis (ISL)  
5 Rue du Général Cassagnou  
68301 Saint Louis  
Frankreich  
www.isl.eu

Universität Bonn  
Institute of Computer Science II  
Friedrich-Ebert-Allee 144  
53113 Bonn  
cg.cs.uni-bonn.de

- [1] Andreas Velten, Thomas Willwacher, Otakrist Gupta, Ashok Veeraraghavan, Mounsi G. Bawendi und Ramesh Raskar, Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging, Nature Communications 3, Article number: 745 (2012), doi:10.1038/ncomms1747
- [2] Adrian Jarabo, Belen Masia, Julio Marco und Diego Gutierrez, Recent advances in transient imaging: A computer graphics and vision perspective, Visual Informatics, dx.doi.org/10.1016/j.visinf.2017.01.008
- [3] Jonathan Klein, Christoph Peters, Jaime Martín, Martin Laurenzis und Matthias B. Hullin, Tracking objects outside the line of sight using 2D intensity images, Scientific Reports 6, Article number: 32491 (2016), doi:10.1038/srep32491