

## Pressemitteilung

Max-Planck-Institut für Intelligente Systeme

Linda Behringer

29.03.2021

<http://idw-online.de/de/news765790>

Buntes aus der Wissenschaft, Personalia  
Gesellschaft, Informationstechnik, Medien- und Kommunikationswissenschaften, Physik / Astronomie, Sprache / Literatur  
überregional



## Die virtuelle Welt noch realistischer darstellen

**Die reale Welt erfassen und Menschen, Objekte und Räume so realistisch modellieren, dass ein Betrachter den Unterschied nicht mehr wahrnimmt – Justus Thies hat sich für die kommenden fünf Jahre ehrgeizige Ziele gesetzt. Die Forschung zu fotorealistischen Videoinhalten bringt eine große Verantwortung mit sich. Die Reinszenierung von Gesichtern zum Beispiel sieht täuschend echt aus; damit wird es immer schwieriger zu erkennen, ob die Person einen Satz auch wirklich gesagt hat. Mit dem Wissen über manipulierte Medieninhalte – sogenannter Deep Fakes – will Thies einen Schwerpunkt auf die Entwicklung von Algorithmen legen, die gefälschte Videoinhalte automatisch erkennen und kennzeichnen.**

Tübingen - Zum 1. April 2021 tritt Justus Thies eine Stelle als Max-Planck-Forschungsgruppenleiter an. Seine Gruppe „Neural Capture & Synthesis“ ist am Tübinger Standort des Max-Planck-Instituts für Intelligente Systeme (MPI-IS) angesiedelt. Thies, der zuvor als Postdoc an der Technischen Universität München (TUM) tätig war, will das Forschungsfeld der Erfassung und Synthese der realen Welt, von Menschen, Objekten, Räumen und sogar ganzen Szenen, maßgeblich prägen. Ziel seiner Forschungstätigkeit ist es, den menschlichen Körper und Objekte so realistisch zu synthetisieren und zu modellieren, dass ein Betrachter den Unterschied nicht mehr wahrnimmt. Basierend auf Deep-Learning-Methoden und dem sogenannten neuronalen Rendering hat Thies' Methode der virtuellen Inszenierung von Gesichtern eine Qualität erreicht, die in Wissenschaft und Medien große Aufmerksamkeit erregt hat. Durch die Kombination neuartiger Methoden des maschinellen Lernens mit klassischen Computergrafik- und Computer-Vision-Ansätzen erzielt Thies vielversprechende Ergebnisse.

Immer realistischere Avatare schaffen

Thies' Forschung könnte in einer Vielzahl von Bereichen Anwendung finden, von autonomen Fahrzeugen oder Haushaltsrobotern bis hin zu medizinischen Anwendungen und der Video-Postproduktion. Sein Hauptziel ist es jedoch, die Telekommunikation zu revolutionieren.

„Ich möchte die Art und Weise verändern, wie Menschen miteinander kommunizieren, die sich nicht persönlich gegenüberstehen können“, sagt er. „Wenn wir in Zukunft mit einer Person sprechen wollen, die sich an einem anderen Ort befindet, werden wir eine Virtual- oder Augmented-Reality-Brille tragen, die einen fotorealistischen 3D-Avatar dieser Person direkt vor uns projiziert. Die Person wird digital so rekonstruiert, dass es sich anfühlt, als wäre sie im selben Raum. Das wird eine erstaunliche Erfahrung sein – allerdings erst dann, wenn das digitale Abbild des Menschen in Aussehen, Mimik und Bewegungen von der realen Person nicht mehr zu unterscheiden ist. Daher ist es mein Ziel, jeden in die Lage zu versetzen, einen 3D-Avatar von sich selbst zu erstellen, der komplett realistisch ist.“

Thies möchte, dass Nutzer sich selbst mit handelsüblichen Kameras erfassen können, ohne auf komplexe Hardware angewiesen zu sein. Menschen zum Beispiel in einem 4D-Körperscanner zu vermessen, der dreidimensionale Scans von Menschen in Bewegung erstellt, wäre viel zu aufwändig, um es in die Breite zu bringen. „Eine Idee ist zum Beispiel, die

Kameras von Spielekonsolen zu nutzen, deren Sensoren Farbe und Tiefe erfassen. Ein eingebauter Abstandssensor misst, wie lange das Licht braucht, um von der Kamera zur Szene und zurück zu gelangen. So erhalte ich eine dreidimensionale Punktwolke einer Person, aus der ich dann einen Avatar erstellen kann," erklärt der Forscher.

Synchronisierung der Bewegung eines Gesichts mit dem Eingangston

Ein weiteres potentiell Anwendungsbereich von Thies' Forschung ist die Video-Postproduktion. In synchronisierten Filmen könnten die Lippenbewegungen der Schauspieler der jeweiligen Sprache angepasst werden. „Um die Mimik und das Aussehen eines bestimmten Gesichts zu lernen, benötigt unser Algorithmus nur einen kurzen Videoclip von etwa zwei Minuten“, sagt Thies.

Diese Technologie, die auch als Echtzeit-Gesichts-Inszenierung, audio-gesteuerte Gesichts-Video-Synthese oder neuronales Puppenspiel bekannt ist, steckt noch in den Kinderschuhen. Sie wird von einem tiefen neuronalen Netzwerk gesteuert, das einen latenten 3D-Gesichtsmodellraum verwendet. Auf diese Weise kann zum Beispiel ein Video erstellt werden, das eine prominente Persönlichkeit zeigt, die mit einer anderen Stimme spricht, oder sogar die Person zeigt, wie sie Dinge mit ihrer eigenen Stimme sagt, die im echten Leben so nie ausgesprochen wurden.

„Das ist die negative Seite des Facial Reenactment, da die Erzeugung von fotorealistischen Bildern missbraucht werden kann. Die Algorithmen, die bei der Lippensynchronisation eines Films angewendet werden, können auch dazu verwendet werden, eine Rede eines Staatsoberhauptes zu kreieren und ihm oder ihr Worte in den Mund zu legen, die gefälscht und unwahr sind. Wenn die Menschen die Fälschung nicht bemerken, könnte dies zu Verwirrung führen, ja sogar eine ganze Gesellschaft untergraben. Im Internet tauchen zunehmend Fälschungen auf, die nur schwer zu erkennen sind. Digitale Multimedia-Forensik ist daher ein weiterer Pfeiler meiner Forschung. Ich möchte Algorithmen entwickeln, die automatisch erkennen, ob eine Videosequenz synthetisch ist oder manipuliert wurde“, so Thies.

Deep-Fake-Technologie verstehen, um sie zu erkennen

Bei der Multimedia-Forensik geht es darum, einen KI-Algorithmus mit einem großen Datensatz zu trainieren, damit gefälschte Videos anhand der in einer Videosequenz sichtbaren Artefakte identifiziert werden können. Solche Artefakte treten oft entlang der Grenzlinie zwischen dem synthetisierten Gesicht und dem realen Hintergrund auf. Mit Hilfe des Algorithmus wird auch das Rauschverhalten der Farben im Bild untersucht, womit es von echten Videos unterschieden werden könnte. Auf diese Weise kann mit Hilfe künstlicher Intelligenz erkannt werden, ob eine Manipulation stattgefunden hat. „Das ist genau die Art von Technik, die dann automatisiert werden kann“, sagt Thies. „Es gibt bereits Versuche, diese Deep-Fake-Erkennungs-Technik als Browser-Plug-In für jedermann anzubieten. So könnten gefälschte Bilder und Videos automatisch erkannt und gekennzeichnet werden. So würden Laien vor Fake News geschützt.“

Die Entwicklung von Multimedia-Forensik ist ein Katz- und Mausspiel, denn Deep-Fakes werden immer besser. Heutzutage ist es relativ einfach, einen Deep Fake zu erkennen. Das wird in naher Zukunft aber immer schwieriger. Die Erstellung fotorealistischerer Videoinhalte ist daher mit einer großen Verantwortung verbunden. „Mit meiner Arbeit lege ich die Messlatte höher, damit nicht jeder einfach ein Video fälschen kann und es unbemerkt bleibt. Durch das Wissen über den Erstellungsprozess können wir die fortschrittlichsten Algorithmen zur Fälschungserkennung entwickeln. Wir müssen bei dieser Technologie ganz vorne mit dabei sein, um die Fälschungen zu erkennen, zu markieren und zu entlarven“, so Thies abschließend.

----

Dr. Justus Thies ist ehemaliger Postdoktorand an der Technischen Universität München. Seit September 2017 arbeitete er im Visual Computing Lab von Prof. Matthias Nießner. Zuvor promovierte er an der Universität Erlangen-Nürnberg, betreut durch Günther Greiner. Seinen Master of Science erhielt Thies von der Universität Erlangen-Nürnberg. Während der Zeit als Doktorand arbeitete er mit anderen Institutionen zusammen, mit Praktika an der Stanford University und

dem Max-Planck-Institut für Informatik in Saarbrücken.

Ein medizinisches Projekt markierte den Beginn von Thies' Forschung zur Erfassung und Synthese von Bild- und Videoinhalten. Ziel war es, die Gesichter von Patienten mit Gaumenspalten zu erfassen. Über mehrere Jahrzehnte wurden die Patienten immer wieder fotografiert. So konnten die Wissenschaftler sehen, wie sich ihre Gesichter im Laufe der Behandlung veränderten und wie ihre Wunden heilten. Mit dem von ihm entwickelten Ansatz konnten neue Patienten mit einer Gaumenspalte sehen wie bestimmte chirurgische Behandlungsmethoden dann später aussehen würden.

Seine Forschung kann auf diesem Youtube-Kanal gesehen werden:  
<https://www.youtube.com/channel/UCwmSTvnV-sjtIINvWYC6Ow>

Pressekontakt:

Linda Behringer  
Communications Manager  
Max-Planck-Institut für Intelligente Systeme  
[linda.behringer@is.mpg.de](mailto:linda.behringer@is.mpg.de)

wissenschaftliche Ansprechpartner:

Dr. Justus Thies  
[justus.thies@is.mpg.de](mailto:justus.thies@is.mpg.de)

URL zur Pressemitteilung: <https://is.mpg.de/de/news/making-the-virtual-world-even-more-realistic>



Dr. Justus Thies  
Justus Thies