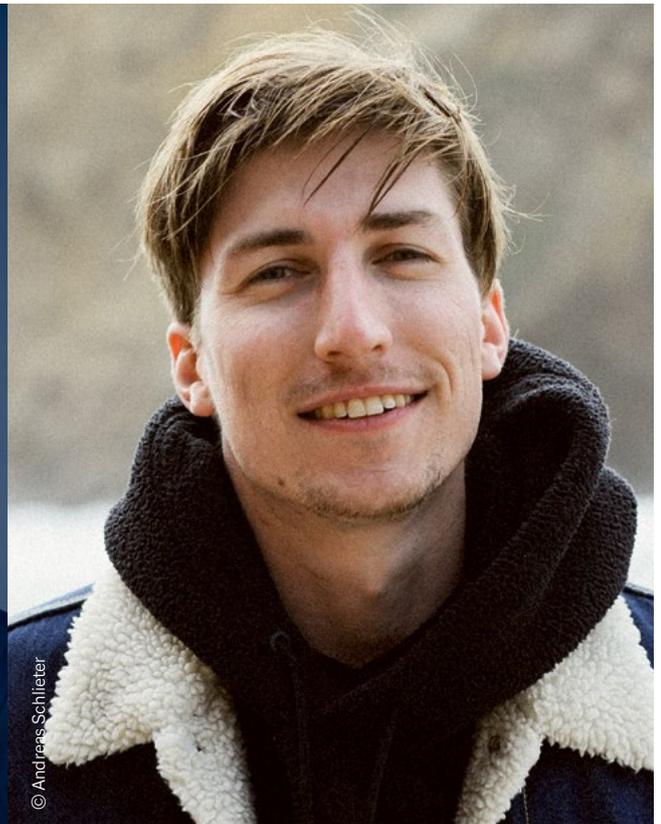


Die KI-Technologie und hier vor allem Generatoren verändern zunehmend auch die Produktion von Filmen, sie revolutionieren Sehgewohnheiten in nie da gewesener Weise. Sven Bliedung von der Heide (SBH), Produktionsstudiobetreiber in Potsdam-Babelsberg, und Pascal Schröder (PS), Produzent und Regisseur aus Hamburg, sprachen mit *mediendiskurs* über Innovationen, Kreativität, das Verhältnis von Mensch und Maschine sowie über die Gefahren von Deepfakes.

Filme



von morgen

Camilla Graubner und Eva Maria Lütticke im Gespräch mit
Sven Bliedung von der Heide und Pascal Schröder

*

Herr Bliedung von der Heide, Sie betreiben auf dem Studio Babelsberg Filmgelände in Potsdam ein volumetrisches Studio. Was können wir uns darunter vorstellen?

SBH: Das volumetrische Capture Studio kann man sich wie einen runden weißen Raum mit Rundumkameras vorstellen. Das Besondere ist: Es gibt nicht eine Hauptkamera, sondern wir erfassen Objekte, Personen, teilweise auch ganze Filmsets komplett in Bewegung, in 3-D.

Man kann dann daraus eine 3-D-Oberfläche erzeugen, die wiederum in Computerspielen, 3-D-Umgebungen oder Filmen eingesetzt werden kann.

Wie viele Kameras braucht es dafür?

SBH: Das ist variabel und kommt auf die Einsatzzwecke an. Wir sehen das eher wie ein System. Teilweise haben wir schon mit acht Kameras gefilmt. Aktuell nutzen wir meistens 42. Damit erreichen wir mit Abstand die höchste Auflösung und sind auf dem Markt weltweit einmalig. Es gibt in Amerika noch ein paar Mitbewerber, die aber alle mit einer Gesamtsystemauflösung von

maximal 500 Megapixel arbeiten. Wir sind bei über 3.000. Das führt im Resultat dazu, dass man bei unseren Figuren auch wirklich die Gesichtsmimik, Falten und Poren sehen kann.

Wenn man sich bei Ihnen abfilmen lässt, entsteht also ein Avatar, der aber noch nicht beliebig in Bewegung versetzt werden kann?

SBH: Ja, wir nennen sie Volumetric Captures oder VoluCaps. Bei sogenannten Motion Capture Sessions werden Personen in Anzüge gesteckt und alle Bewegungsmuster, die benötigt werden, aufgenommen. Anschließend wird ein Avatar animiert. Mit Volumetric Capture hingegen wird die echte Person mit ihren echten Bewegungen und Emotionen direkt aufgenommen - diese können also nicht mehr geändert werden. Das ist wie eine 2-D-Aufnahme, nur in 3-D. In Zukunft sollen jedoch auch neue Animationen und Bewegungen möglich sein. Wir arbeiten gerade im Rahmen eines Forschungsprojekts mit dem Fraunhofer-Institut daran. Im Augenblick ist es noch problematisch, neue emotionale

Gesichtsausdrücke zu erzeugen oder beispielsweise Kleidung, die bei Bewegungen immer wieder anders fällt, natürlich aussehen zu lassen. Unser Fokus liegt darauf, dass die Authentizität der Person hundertprozentig gegeben ist und man nicht das Gefühl hat, es wäre eine Art Computerspiel-Avatar, der animiert wurde. Das kriegt man momentan nur hin, wenn man die Personen in Bewegung und Mimik eins zu eins aufgenommen hat.

Derzeit arbeiten Sie am Film *Mickey17*. Sie dürfen wahrscheinlich nicht viel verraten, aber könnten Sie an diesem Beispiel einmal erklären, wo da die Ersparnisse sind?

SBH: Bei *Mickey17* haben wir viele Sachen gemacht, die es so vorher noch gar nicht gab. Man kann da also im Moment schwer einen Vergleich ziehen. Eine klassische Kostenreferenz ist, wenn ich eine Person in eine 3-D-Umgebung stellen möchte, beispielsweise in einem Computerspiel, einer Virtual-Reality-Umgebung oder in Augmented Reality, die dann etwas erzählt und möglichst echt aussehen soll. Die Erstellung einer solchen Computerspielfigur liegt im Bereich von 5 bis 6 Mio. Euro. On top sind die Kosten für das Animieren und das Integrieren in die Spielumgebung. In unserem Studio

brauche ich wenig Zeit für das Aufnehmen des Materials in 3-D. Wenige Tage später nach dem Drehtag ist es fertiggerechnet. Das andere Verfahren dauert Monate. Mit dem Drehtag und allem Drumherum sind das bei uns vielleicht Kosten von 10.000 Euro. Vergleichsweise ein Witz, wenn man bedenkt, dass z.B. die ABBA-Show, bei der die Mitglieder der Band als digitale Avatare aufgetreten sind, ungefähr 170 Mio. Euro gekostet hat.

An dieser Stelle möchte ich Herrn Schröder in die Runde holen. Ihre Firma wurde kürzlich umbenannt in FLAPPERS FILM & VFX. Sie produzieren Filme...

PS: ... und visuelle Effekte.

Was heißt das genau?

PS: Hinter „visuelle Effekte“ verbirgt sich eine Disziplin, die sowohl in Dokumentationen als auch in Spielfilmen eingesetzt wird. Wir werden vor allem in der Postproduktion angefragt. Es geht darum, dass das, was man im Bild nicht live herstellen kann, nachträglich eingefügt wird. Als Beispiel: Wir hatten eine Filmproduktionsfirma, die einen Film gedreht hat und ein falsches 3-D-Hotelschild im Bild hatte. Unsere Aufgabe war es, das Schild nachträglich auszutauschen. Konkret bedeutet das, verschiedene Arbeitsschritte zu durchlaufen, in diesem Fall: Drohnenfahrten um das 3-D-Schild herum, das muss getrackt werden, das Schild nachgebaut und an die Lichtverhältnisse angepasst werden. Teilweise können diese Schritte mit KI umgesetzt werden. Beim

Tracking bewegt sich die Kamera z.B. um Trackingpunkte herum, die der Kamerabewegung eine Orientierung geben. Bei einer „Fahrt“, die kompliziert ist, dauert das manchmal für eine 3-D-Einstellung drei Tage. In Zukunft wird es möglich sein, selbst komplizierte Aufgaben innerhalb weniger Klicks zu tracken. Ziel sollte dabei sein, keinen qualitativen Unterschied zu erkennen – dann haben wir unsere Arbeit gut gemacht.

Ist der Text-zu-Video-Generator Sora aus Ihrer Sicht vielversprechend?

SBH: Absolut.

PS: Je nachdem für was, glaube ich. Wenn man einen Spielfilm machen möchte und Konsistenz erreichen will, dann sehe ich für die nächsten zwei Jahre noch Schwierigkeiten.

SBH: Ja, Konsistenz ist so ein Thema, wobei das immer besser wird. Auch die Directability, also die Einflussnahme durch die Regie, wird immer besser. Das Problem bei Sora und OpenAI generell ist, dass man sehr limitiert ist. OpenAI hat ein sehr großes „Ethik-Department“, was natürlich für „normale“ Menschen, die Anfragen stellen, super ist. Doch wenn man im kreativen Prozess Freiheiten braucht, ist das ein Problem, beispielsweise beim Drehbuchschreiben. Man hat so viele Limitationen und Filter drin, dass es oftmals einfach schwierig ist, sich dort überhaupt kreativ auszuleben und auch kontroverse Geschichten zu erzählen, weil vieles weggefiltert wird. Ähnlich ist es bei den Video- und Bild-Generatoren von OpenAI. Will man interessantere Geschichten erzählen oder Bilder schaffen, die auch teilweise wichtig für Filme sind, stößt man oft an Grenzen, weil es sich aufgrund ethischer Filterungen nicht erzeugen lässt. Das sehe ich als großen Nachteil. Gefühlt ist man vielleicht ein Dreivierteljahr davon entfernt, bis man so etwas Ähnliches dann auch im Open-Source-Bereich haben wird. Wenn die Einschränkungen damit wegfallen, wird das Thema richtig stark aufleben.

„Früher hätte man eine Welt nachbauen müssen.“

Pascal Schröder

Nutzen Sie bei Ihrer Arbeit derartige Generatoren?

PS: Ja, sogar relativ viel. Wir nutzen z.B. Matte Paintings. Für eine Dokumentation sollte ein historisches Hamburg erzählt werden. Dafür haben wir ein Original-Schwarz-Weiß-Bild von 1920 eingespeist. Das hatte von der Auflösung her eine relativ schlechte Qualität. Wir haben das dann mit einer AI hochskaliert und in Farbe umgewandelt. Ergänzt mit den ganz klassischen VFX-Diszi-

plinen – Unschärfe reinbringen, Staub hinzufügen, noch ein paar Bewegungen wie Vögel ergänzen – sieht das Foto am Ende sehr, sehr echt aus. Früher hätte man eine Welt nachbauen müssen. Und jetzt war das von der Recherche bis zum Endergebnis innerhalb von zwei Stunden erledigt.

War das Ergebnis auch zufriedenstellend?

PS: Ja, für das Fernsehen reicht es tatsächlich, bei einem Kinofilm eher nicht. Wir wollten z.B. über das Ruhrgebiet etwas machen, also das Ruhrgebiet der 1970er-Jahre erzählen. Konkrete Daten dafür sind im Internet kaum vorhanden. Das ist dann schwierig zu generieren. Wenn man aber ein jetziges Hamburg erzählen möchte,

dann ist es einfach. Es ist also immer eine Frage der Korrelation: Was möchte man? Welche Daten sind vorhanden? Und welche Umwege gibt es, um zum Ergebnis zu kommen?

Herr Bliedung von der Heide, Sie erwähnten einmal, dass die Daten, die zum Trainieren der KIs verwendet werden können, „so langsam endlich sind“. Was genau haben Sie damit gemeint?

SBH: Bevor es ChatGPT gab, konnte man einfach alle Texte aus dem Internet nehmen und wusste, dass die hauptsächlich von Menschen geschrieben sind. Seit ChatGPT ist das anders, da ist ein Großteil der Texte nicht mehr von Menschen geschrieben. Um solch eine AI aber zu verbessern, muss sie mit Daten, die nicht aus einer AI stammen, gefüttert werden. Sonst entwickelt sich da nichts weiter. Bei Bildern ist es das gleiche Problem: Je mehr Bilder nur noch KI-generiert sind, desto schwieriger wird es, die Daten zu erweitern. Wie macht man also weiter? Wie kommt man an neue Daten?

Aktuell bekommen wir bereits Anfragen, dass wir Daten aufnehmen, die für solche Trainingssysteme nutzbar sind. Wir gehen davon aus, dass bereits in einem Jahr massiv viele Fernsehbeiträge oder Nachrichten KI-generiert sein werden oder in zwei Jahren ganze Kinofilme. Dann ist die entscheidende Frage: Auf welchen Trainingsdaten kann man nachher neue KI-Systeme trainieren, wenn fast alles maßgeblich mit KI erzeugt worden ist?

Dazu passt das Thema „Fake News“, das mit den KI-Generatoren auf ein nächstes Level gehoben wird. Gibt es Ideen, beispielsweise Wasserzeichen, um Inhalte zu kennzeichnen?

SBH: Wir arbeiten tatsächlich daran, eine Erkennung zu schaffen. Wir denken an eine Plattform, die es ermöglichen soll, für Nachrichtensender und auch andere Dienste zu prüfen, ob das Material ein Video-Fake ist bzw. mit KI generiert wurde. Denn noch erzeugen diese ganzen Systeme gewisse Artefakte, beispielsweise in Schattierungen oder Knochenabständen bei der Gesichtsmimik. Wir trainieren auf Grundlage der Artefakte, um diese erkennbar zu machen und so sehen zu können, wenn etwas inkonsistent, also Fake ist. Es gibt nach meinem Kenntnisstand in Europa noch niemanden, der wirklich an einem brauchbaren Verfahren arbeitet.

PS: Ich habe da auch ein Beispiel. Vor zwei Monaten habe ich eine neue KI ausprobiert. Oft sind die KIs über das Programm Discord zugänglich. Dort kann jeder prompten - und andere Leute können zuschauen, was erstellt wird. Ich konnte live mitverfolgen, wie jemand versucht hat, von Selenskyj ein Deepfake-Video zu machen. Die Qualität wurde mit jedem Versuch besser! Es ist mittlerweile so einfach, diese Videos herzustellen. Das kann wirklich jeder, da muss man gar kein Tekkie sein.

Sie beschäftigen sogar Mitarbeiter, die sich ausschließlich mit den KI-Entwicklungen auseinandersetzen und diese ausprobieren, oder?

PS: Ja, das ist richtig. Um optimale Ergebnisse zu erzielen, muss man AIs mit Daten trainieren. Das ist z.T. auch eine Fleißarbeit. Es gibt unglaublich viele Tools, die man miteinander verbinden muss, um voranzukommen. Wir sehen uns als Pioniere, die auf dem Stand sein wollen, um KI auch für Spielfilme nutzen zu können. Es gibt bereits viele Kurzfilme im Internet, die eine gewisse Qualität haben, aber das reicht noch nicht aus, um Spielfilme damit zu produzieren. Unser Ziel ist es, die Ersten in Deutschland zu sein, die dies wirklich nutzbar machen können und dabei eine Qualität erreichen, die für Spielfilme geeignet ist. Dafür müssen wir verschiedene AIs

miteinander verbinden und forschen, was Sinn macht. Es bringt nichts, ein Technikfeuerwerk zu entfachen, wenn am Ende die Geschichte nicht dazu passt. Je mehr KI zur Verfügung steht, desto klarer muss die Vision eines Kreativen, einer Regisseurin, eines Drehbuchautors sein. Die richtigen Entscheidungen müssen getroffen werden, um eine klare Vision zu verwirklichen.

SBH: Du meinst, dass auch der Regisseur oder Produzent im klassischen Sinne noch gebraucht wird? Dass weiterhin Menschen nötig sind, die kreative Entscheidungen treffen?



© Volucap_Studio



© FLAPPERS FILM & VFX

PS: Ja, das glaube ich tatsächlich. Man sieht es z.B. bei Netflix oder anderen großen Playern, die auf Parameter wie IPs (Intellectual Properties) setzen. In Zukunft wird es noch datengetriebener sein. Aber am Ende geht es ja auch immer noch darum, dass eine Person eine Entscheidung treffen muss, um vielleicht etwas komplett Einzigartiges zu machen, etwas, bei dem menschliche Fehler enthalten sind, damit es etwas Besonderes wird. Marketing spielt auch eine große Rolle. Vielleicht gibt es auch in Zukunft menschengemachte Filme als weiteres Genre – und das ist dann das neue Arthousekino.

„Es wird natürlich auch Menschen geben, die niemals einen KI-Film schauen werden.“

Sven Bliedung von der Heide

Linke Seite:

Oben: Blick in das Volucap Studio

Unten: Ein Ausschnitt aus der Previsualisierung eines Filmprojekts, in dem eine Traumwelt erschaffen werden sollte. Das Bild wurde vollständig von AI generiert.

SBH: Ich würde dagegen argumentieren, dass wir Menschen nicht gerade sehr gut darin sind, Filme zu machen. Wir haben einen enormen Output an Material, oft scheitern sogar aufwendig produzierte Filme, die Millionen kosten, daran, ihr Publikum zu erreichen. Gerade wenn es um die Schaffung von Individualität in Filmen geht, ist das ein Dilemma. Wir wollen Filme produzieren, die viele Menschen gleichzeitig erreichen, um profitabel zu sein. Mit KI haben wir jedoch die Möglichkeit, viel kreativere Filme zu schaffen, die auf den Einzelnen zugeschnitten sind und nicht so stark dem Mainstream entsprechen. Die Mainstream-Inhalte werden wahrscheinlich zuerst von KI erstellt werden, da sie auf das trainiert wird, was die meisten Menschen mögen. Das wird vermutlich der schnellste Fortschritt sein, den wir mit KI erzielen können. Schwieriger wird es bei Arthousefilmen und ähnlichen Projekten sein. Es wird natürlich auch Menschen geben, die niemals einen KI-Film schauen werden, genauso wie es Menschen gibt, die sich weigern, Netflix zu nutzen. Ich vermute, dass wir in Zukunft die Möglichkeit haben werden, individuell Filme on the fly zu generieren. Die KI lernt ständig von den Daten, was mir gefällt und was nicht, auch in Bezug auf die Dramaturgie. Ich glaube, die KI wird uns in diesem Bereich übertreffen. Es wird schwierig sein für Menschen, auf individueller Ebene mit der KI zu konkurrieren, wenn es um die Erstellung oder Entscheidung von Inhalten geht. Die Frage ist eher, wie die großen Plattformen untereinander konkurrieren werden, wenn alle dieselbe KI nutzen.

Wo sehen Sie die Grenzen des Möglichen?

PS: Die Persönlichkeitsrechte müssen gewahrt werden. Man darf die Rechte anderer nicht mit Füßen treten, nur weil es möglich ist.

SBH: Was man aus deutscher und europäischer Sicht sagen muss, ist, dass wir generell ein sehr starkes Persönlichkeitsschutzrecht haben. Das ist nicht selbstverständlich! Das, was in Amerika beispielsweise durch Schauspielstreiks erreicht wurde, gilt bei uns sowieso schon per Gesetz. Wir haben eine solide Basis in dieser Hinsicht. Nehme ich z.B. das Gesicht einer Person für Deepfakes und das kommt heraus, kann ich dafür verklagt oder abgemahnt werden, weil es gegen das Gesetz verstößt. Die Frage ist jedoch, ob weitere Einschränkungen notwendig sind, wenn es um die Nutzung von Deepfakes oder KI geht, zu denen wir bereits bestehende Persönlichkeits- und Schutzrechte haben. Das wird etwas schwieriger, denn oft zielt man eher darauf ab, Einschränkungen bei Trainingsdaten oder Ähnlichem vorzunehmen. Wenn Europa dies teilweise einschränken würde, liefe man Gefahr, im Vergleich mit dem Ausland nicht mehr auf dem neuesten Stand zu sein. Es ist also ein Drahtseilakt. Das Problem ist auch, dass KI noch so

neu ist, dass es sinnvoller ist, flexible Modelle zu haben, anstatt einen festen Katalog mit Richtlinien zu erstellen, da sich die Möglichkeiten und Erwartungen alle paar Wochen ändern. Die ethischen Aspekte sind bereits durch die bestehenden Gesetze geregelt. Die wirkliche Herausforderung sehe ich in den Bereichen außerhalb des normalen Rechtsrahmens. Wenn beispielsweise eine russische Gruppe Deepfakes erstellt und in Telegram-Chats hochlädt, gibt es keine Instanz, die das verfolgen kann.

Was außerdem noch spannend ist, ist die Frage des Urheberrechts, insbesondere im Zusammenhang mit der Generierung von KI-Videos. Das könnte ein interessantes Alleinstellungsmerkmal oder ein Unique Selling Point für euch VFX-Künstler sein, denn das Problem ist, dass komplett von KI generierte Inhalte derzeit nicht geschützt sind. Wenn Sora mir ein Video ausspuckt und ich es in meinem Film verwende, ist so ein Prompt per se nicht schützbar. Jeder könnte dann dieses Stück aus meinem Film verwenden, da kein menschliches Handeln in der Geschichte involviert war.