2020
3
ISSN 1433-2620 > 24. Jahrgang >> www.digitalproduction.com
Publiziert von Pixeltown GmbH

Deutschland € 17,90
Österreich € 19,–
Schweiz sfr 23,–

DIGITAL PRODUCTION

DIGITAL PRODUCTION

MAGAZIN FÜR DIGITALE MEDIENPRODUKTION

MAI | JUNI 03:2020

Ausbildung
Was macht eine gute VFX-Uni aus?

Tests
Eizo, Philips, Soundweaver, Woosh und Akeytsu

Projekte
Mulm, The Witcher, Marvel Heros, Walking Dead: Maya

und vieles mehr
Flame, Blender, InstaLOD, Nuke 12.1, Topaz und mehr

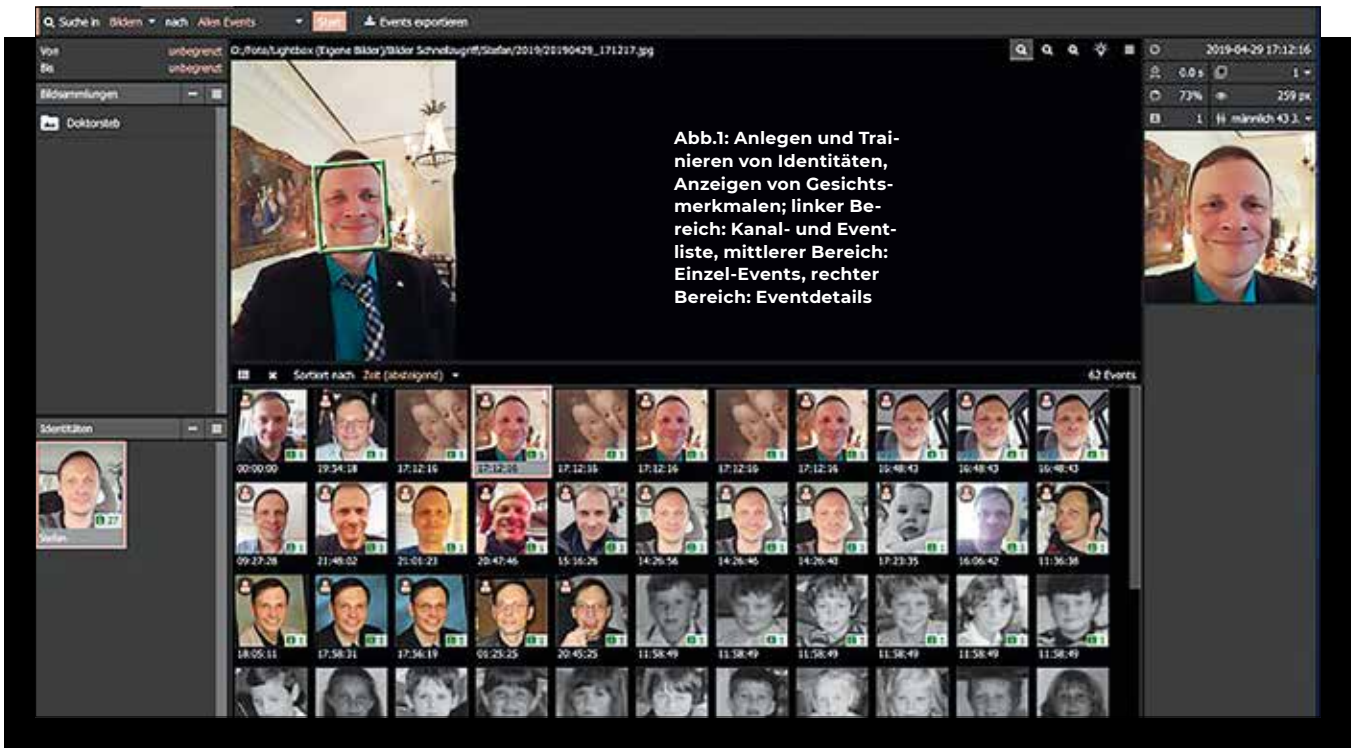

Abb.1: Anlegen und Trainieren von Identitäten, Anzeigen von Gesichtsmerkmalen; linker Bereich: Kanal- und Eventliste, mittlerer Bereich: Einzel-Events, rechter Bereich: Eventdetails

Gesichtserkennung in der Forensik

Im Nachgang unserer KI-Ausgabe kamen ein paar Leserfragen: Was ist denn das mit der KI? Worauf schaut die bei Gesichtern? Und wann kann ich meine Tools zum Finden bestimmter Gesichter in meinem Footage verwenden? Und da uns außer dem berühmten XKCD (**xkcd.com/2173/**) nichts eingefallen ist, haben wir einen offiziellen Forensiker gefragt, was es denn damit auf sich hat. Und wo ist der Unterschied zur Forensik, in der Gesichtsanalyse einen festen Platz hat?

Ganz allgemein gesagt ist die Gesichtserkennung (Face Recognition) ein biometrisches, KI-gestütztes und automatisiertes Diagnoseverfahren, mit dem die Zuordnung zu einer Person gelingt. Die Gesichtserkennung basiert dabei auf mathematischen Verfahren zur Mustererkennung und maschinellem Lernen. Maschinelle Lernverfahren (Deep Learning), basierend auf Data Mining (Finden neuer Muster und Gesetzmäßigkeiten), generieren Wissen aus Erfahrung. Künstliche Systeme lernen aus Beispielen und verallgemeinern die Erkenntnisse nach Beendigung der Lernphase. Erste Ansätze der Gesichtserkennung gab es bereits 1958 mit nicht lernenden künstlichen neuronalen Netzwerken (KNN).

Personensuche

Wird von der Zuordnung der Person gesprochen, handelt es sich im forensischen Sinne um eine Identifikation und Verifikation, um eine Authentifizierung von Personen. Im Gegensatz dazu ist die Gesichtsdetektion abzugrenzen, die nur eine Lokalisierung eines Gesichts im Bild/Videobild vornimmt. Gesichtsdetektionen sind bereits flächendeckend in Bild- und Videokameras und mobilen Endgeräten im Einsatz. Von der professionellen Gesichtserkennung abzugrenzen sind auch als Apps erhältliche Werkzeuge, mit denen Gesichter erkannt und verändert werden können. Diese dienen spielerischen, benutzeridentifikationsbezogenen bzw. marketingstrategischen Zwecken im Social-Media-Umfeld für gezielte Werbung im Internet.

Während die Fehlerrate bei zweidimensionalen (2D) Verfahren (Gesichtserkennung in Bildern) 1993 noch bei 79 % lag, können seit 2006 Fehlerraten unter 1 % erreicht werden. Vor allem bei den dreidimensionalen (3D) Verfahren der Echtzeit-Videoanalyse (Gesichtserkennung aus Videos) werden heute hohe Erkennungsgenauigkeiten erreicht, die unabhängig von (Gesichts-)Posen sind.

Eckdaten

Zur Klassifikation zählen die Identifikation sowie die Geschlechts- und Altersschätzung. Der minimale Augenabstand, ab dem ein Gesicht klassifiziert wird, beträgt 12 Pixel. (Gesichts-)Templates enthalten die Ergebnisse der Gesichtsanalyse von stark sichtbaren Merkmalen des frontalen Kopfes, mit der eine Identifikation möglich wird. Gesichts-Templates werden nur dann zur Identifikation herangezogen, wenn sie eine bestimmte Mindestqualität erreichen (mind. 50 % Ähnlichkeit). Gedrehte, verdeckte und unscharfe Gesichter haben typischerweise eine geringe Gesichtsqualität, wodurch sie von vorneherein von der Analyse ausgeschlossen werden sollten.

Aus auf 2D basierenden Bildmaterialien kann ein sogenanntes Template ermittelt werden. Videomaterialien ermöglichen, je nach Inhalt und Laufzeit, ein Vielfaches an Templates. Es werden geometrische Anordnungen analysiert, Textureigenschaften der Oberfläche festgestellt und Identitäten (basierend auf Trainingsdaten) ermittelt. Trainingsdaten sind nicht nur von der Maschine auswendig gelernte Beispiele, sondern aufgebaute Algorithmen, aus denen ein statistisches Modell entsteht und Muster sowie Gesetzmäßigkeiten in den Trainingsdaten erkannt werden können. Damit wird es möglich, unbekannte Daten (Gesichter in neuen Bildern und Videos) zu erkennen und diese als Identität zu ermitteln und abzuspeichern.

Bevor wir tiefer einsteigen: Ich habe bewusst die Herstellerbezeichnungen weggelassen. Die Softwareanwendungen sind

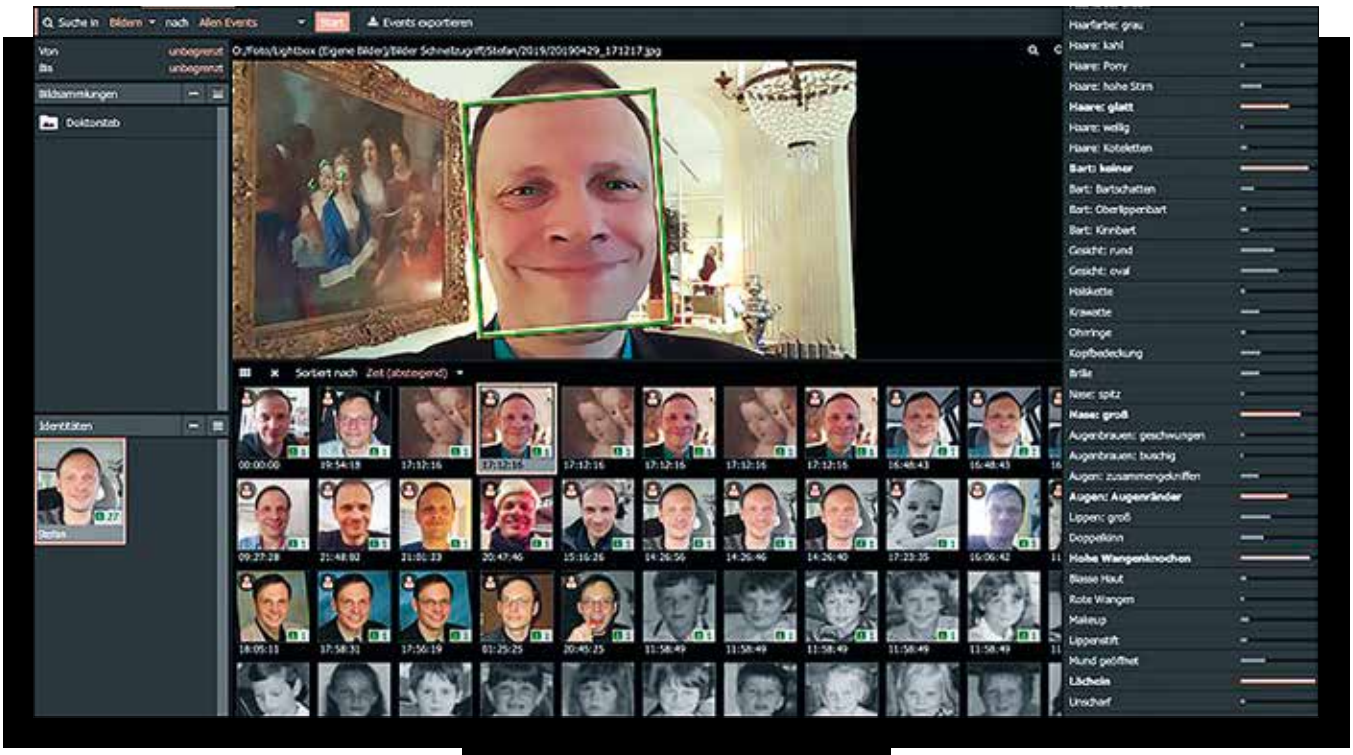

Abb. 2: Verwalten von Identitäten, Anzeigen von Gesichtsmerkmalen

komplexe Programmierungen, die in Verbindungen mit bestimmten Computerdiensten und eigens dafür installierten Datenbanken auf Hochleistungsrechnern laufen. Mit einem normalen Rechner ist es meist nicht möglich, die Software einzusetzen. Ich kenne inzwischen 6 verschiedene professionelle Gesichtserkennungssysteme, es gibt sicherlich noch einige mehr, die für forensische Analysen eingesetzt werden können. Einige dieser Systeme werden z.B. von Firmen wie Facebook, Instagram, WhatsApp und Google genutzt. Diese Firmen halten sich über die eingesetzten Algorithmen allerdings sehr bedeckt.

Im Bild zu sehen ist der Arbeitsdesktop der forensischen Anwendung, der in verschiedene Verwaltungsbereiche aufgeteilt ist.

Die Kenntnis der normalen Erbmerkmale findet im außermedizinischen Bereich ihre Anwendung. Häufig werden in forensischen (gerichtlichen) Gutachten für ein Erkennen und Zuordnen von Personen biologisch-anthropologische Vergleiche anhand von Nasen-, Augenbrauen- und Ohrenformen zur Klärung von Abstammungs- und Identitätsfragen durchgeführt. Während die Abstammungsprüfung (Klärung strittiger Verwandtschaft) bei der Gesichtserkennung nicht von Bedeutung ist, ist der Identitätsprüfung, der Frage der Identität einer Person in Form von Identitätsnachweis oder -ausschluss, eine besondere Aufmerksamkeit zu schenken. Der Rückschluss auf Identität bleibt eine Wahrscheinlichkeitsaussage, da der Identitätsnachweis eine Vielzahl übereinstimmender Merkmale verlangt (Problematik eineiige Zwillinge, Doppelgänger). Diese hängen von der Merkmalsdiagnose und der Häufigkeit der jeweiligen Merkmalsausprägung in der Bevölkerung ab. Der Identitätsausschluss ist eine absolute Aussage. Eine einzige eindeutige Differenz reicht aus, die Identitätsfrage zu verneinen. In der Regel läuft es darauf hinaus, ob zwei Fotos oder Fotosätze oder Videos mit den fraglichen Personen dieselbe Person darstellen.

Bei einer Foto-Identifikation handelt es sich meist um Aufnahmen aus Raumüberwachungen (Diebstahl, Bankraub etc.), einer Verkehrskontrolle, vermissten Personen oder historischen Aufnahmen von Personen mit großen Zeitabständen.

Merkmale

Das biometrische, KI-gestützte, automatisierte Diagnoseverfahren zur Gesichtserkennung ist in der Lage, sehr viele Gesichtsmerkmale gleichzeitig zu erfassen und einer Identität zuzuordnen bzw. diese Identität zu erstellen, anhand derer jederzeit weitere Merkmale trainiert werden können (dynamische Erweiterung).

Das Erstellen von Identitäten basiert auf der Feststellung von Gesichtsmerkmalen. Diese Gesichtsmerkmale können sehr vielfältig vorhanden sein. Zu den typischen Gesichtsmerkmalen zählen in der retrospektiven Video- und Bildanalyse u.a. Alter, Geschlecht, Augen, Augenringe, Ohren, Nase, Mund, Augenbrauen, Stirn, Kinn, Philtrum, Bart, Brille, Ohrringe, Mimik (z.B. Lächeln), Lippen, Lippenstift. Bei der Analyse werden Dutzende dieser Merkmale automatisch lokalisiert.

Insbesondere Augenabstand (optimal größer als 50 Pixel), Kopfdrehung und Kopfwinkel dürfen eine bestimmte Mindestanforderung nicht unterschreiten. Entscheidend sind gute Materialien, die eine homogene Ausleuchtung (keine Schattenwürfe im Gesicht, keine Über- oder Unterbelichtungen) und eine gute Bildqualität (keine Unschärfen, Interlacing-Effekte, Kompressionsartefakte) bieten. Je mehr Bild- und Videomaterialien von einer Person vorliegen, desto besser kann eine Identität trainiert werden (optimal mindestens 100 Einzelbilder von möglichst vielen unterschiedlichen Events).

Die Gesichtserkennung ist insbesondere bei Social-Media-Plattformen durch ein Anlegen von Identitäten kritisch zu betrachten. Werden die vom Nutzer zuerst hochgeladenen Bilder später gelöscht, ist fraglich, was genau denn gelöscht wird. Im schlechtesten Fall wird nur der Zugriff auf das Bild entfernt. Das Bild selbst und die Identität bleiben erhalten. So ist es jederzeit möglich, zu einem späteren Zeitpunkt neu hochgeladene Bilder sofort wieder der entsprechenden Person zuzuordnen.

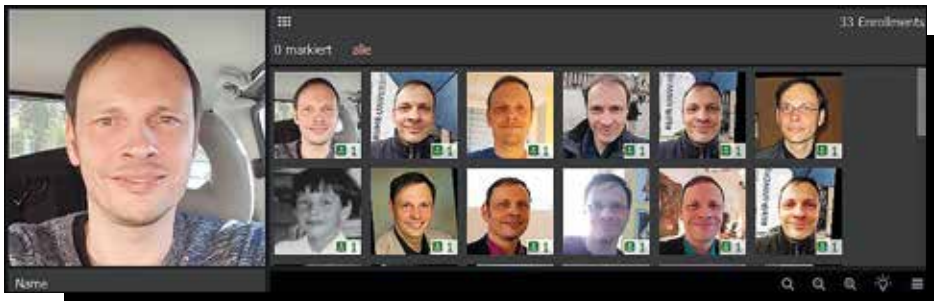

Abb 3: Enrollments (Trainingsmaterial) nach der Gesichtsanalyse, Anzahl festgestellter Templates

Abb 4: Verwalten von Trainingsmaterial, Hinzufügen zu Identitäten (Identitäten trainieren)

Abb 5: Erkennen und Trainieren von Identitäten; linker Bereich: Suchlisten für Medien und Identitäten; mittlerer Bereich: Medienplayer und Suchergebnisse; rechter Bereich: Eventdetails

Abb 6: Festgestellte Identität und Anzahl der vorhandenen Templates

Abb 7: Ähnlichkeiten passender Events

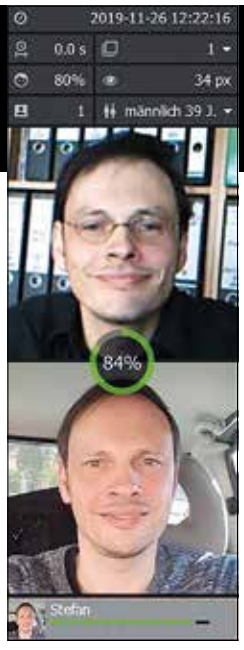

Abb 8: Eventdetails: Grad der Ähnlichkeit einer Identität (Identitätsvergleich)

Abb 9: Gesichtslandmarken

Hier müsste jemand z.B. Google verklagen, wenn er weiß, dass seine Bilder trotz Löschung noch vorhanden sind. Allein diese Hürde ist schon sehr hoch, überhaupt an die Information zu kommen. Dann müsste noch der Beweis hierfür erbracht werden. Deshalb ist unser Digitalzeitalter sehr kritisch und mit Vorsicht zu betrachten. Jeder ist für seine Daten selbst verantwortlich. Was einmal hochgeladen ist, ist für immer irgendwo gespeichert, egal was danach passiert. Bereits beim Upload findet der Scan, also die Gesichtserkennung statt. Der Upload kann online (z.B. bei Instagram etc.) oder auch offline (in die eigene Analysesoftware) erfolgen.

Den Gesichtsmerkmalen werden sogenannte Gesichtslandmarken, insbesondere die Frontallandmarken eines Gesichts zugeordnet. Perspektivische Objektdarstellungen haben Einfluss auf den Grad einer Übereinstimmung von Gesichtslandmarken, die im Zentrum eines Bildes genauer sind als an deren Rändern. Mithilfe dieser Landmarken und eines adaptiven 3D-Gesichtsmodells können Gesichtsform und Identität bestimmt und zugeordnet werden.

Adaptive 3D-Modelle?

In erster Linie sind die Gesichtserkennungssysteme auf die Analyse ausgelegt und nicht auf eine erneute Ausgabe – ob ein Zugriff möglich ist, hängt immer davon ab, ob der Hersteller seine Software überhaupt so konzipiert hat, dass solche Daten erfasst werden können. Für diese Disziplin gibt es bereits andere Anwendungen, die in der Film- und Musikindustrie eingesetzt werden. So gehen bereits verstorbene Sänger und Sängerinnen wieder auf Tour bzw. geben Konzerte (Whitney Houston, Michael Jackson, Roy Orbison, Abba). Aber auch verstorbene Schauspieler und Schauspielerinnen erfahren ihre Wiederauferstehung. 2020 soll ein neuer Film („Finding Jack") mit James Dean (1955 bei einem Autounfall verunglückt) in die Kinos kommen. Auch Carrie Fischer war längst tot, als sie noch einmal in die Rolle der Prinzessin Leia („Star Wars") schlüpfte. CGI bedient sich übrigens in gewisser Weise auch der Deep Fakes – als Grundlage dienen alte Filme, Bilder und anderes Archivmaterial. Digitale Schauspieler sind die Zukunft. Und sicherlich werden auch Filmstudios neue (digitale) Schauspieler kreieren.

Probleme bei der Gesichtserkennung können Merkmalsveränderungen bei Personen sein. Solche Veränderungen entstehen bei sehr großen zeitlichen Distanzen zwischen den Vergleichsobjekten, aber auch bei Unfällen, chirurgischen Eingriffen, hormonellen Störungen oder Alkoholsucht.

Veränderungen treten auch bei der Objekterfassung auf. So lassen sich Veränderungen durch unterschiedliche Körperhaltung oder Mimik herbeiführen. Viele Eigenschaften der Objektabbildung stellen kein Problem mehr dar – heutige 3D-Analysen erfassen sehr viele Gesichtsmerkmale. Allerdings können immer noch fototechnische Umstände (Beleuchtung, Schärfe,

Auflösung etc.) Differenzen oder nicht vorhandene Übereinstimmungen vortäuschen. Auch bei zweidimensionalen Objektabbildungen gibt es Probleme, da sie eine Profilansicht des Gesichts vermissen lassen.

Deep Fakes

Schwachpunkte bei der Gesichtserkennung sind sogenannte Deep Fakes (auch Deepfake oder DF). Mithilfe von KI werden täuschend echt wirkende Bilder und Videos erstellt, die jedoch nicht echt sind (gefälschte Medien). Deep Fakes basieren auf KNN, d.h. es findet eine Abstraktion in der Informationsverarbeitung statt, die gefälschten Medien entstehen weitgehend autonom. Für ein Video-DF reichen bereits 15 Sekunden Quellmaterial einer Person für deren Antizipation, also Vorwegnahme eines zukünftigen Verhaltens von Gesichtsbewegungen aus.

Hierzu werden 3 DF-Techniken eingesetzt: „face swap", „lip-sync" und „Puppet-Master". Derzeitige DF-Videos zeigen Schwächen, anhand derer DFs (noch) erkannt werden können. Z.B. zeigt der Tausch eines Gesichts (face swap) oftmals Unregelmäßigkeiten im Umfeld und Hintergrund der manipulierten Stellen. Personen sitzen oder stehen deshalb oft auffällig ruhig. Auch die Lippensynchronität (lip-sync) lässt zu wünschen übrig. Es entstehen Missverhältnisse zwischen Audiospur und Lippenbewegungen (Asynchronität). Oft werden nur die Lippenbewegungen verändert, nicht jedoch der Gesichtsausdruck.

Und für VFX?

Die meisten Parodisten setzen hier auf Übertreibungen. Das gilt sowohl für Mimik und Bewegung sowie Sprache als auch für Comicillustrationen. Durch die Übertreibung wird der Charakter sofort erkennbar. Bei Deep Fake geht es um Imitation. Das ist schon deutlich schwieriger. Der Charakter muss ausführlich studiert werden. Vieles in unserem Verhalten läuft unbewusst ab. Ich sehe es als sehr schwierig an, viele unserer unbewussten Bewegungen und Äußerungen, derer wir uns selbst kaum bewusst sind, als Imitator vorherzusehen. In der Forensik beschäftige ich mich ebenso mit der Sprecherverifikation. Hier habe ich intensive Untersuchungen durchgeführt, ob z.B. Dialekte, Akzente und Fremdsprachen (nicht Muttersprache) fehlerfrei imitiert werden können. Um die Antwort vorwegzunehmen, es ist sehr schwierig, um nicht zu sagen unmöglich.

Die meisten Deep-Fake-Videos, die im Netz kursieren, besitzen (noch) nicht die Qualität, um nicht als Fake erkannt zu werden, z.B. Merkel/Trump-Morphings, Trumps Fake-Video „Aids is over" oder das Fake-Video mit Barak Obama von Jordan Peele. Aber um es ganz klar zu sagen: Es gibt wirklich hervorragende Fake-Produktionen. Da ist es so ohne Weiteres nicht mehr möglich, Ansatzpunkte zu finden, die nicht stimmig sind. Der Fake wird in der Tat nur dann aufgedeckt, wenn sich der Produzent des Videos outet. Solche Videos werden z.B. als verdecktes Marketing eingesetzt, um Produkte zu bewerben, aber auch viral laufen solche Videos oft sehr erfolgreich. Für das ungeübte Auge ist der Unterschied nicht mehr feststellbar. Auf n-tv laufen tolle Dokumentationen, die sich mit der Frage Fake/Not Fake auseinandersetzen.

Es ist nicht die Frage, ob es möglich ist, Stars wiederzubeleben. Es gibt genügend Beispiele, dass dies perfekt gelingt. Eine Unterhaltungsindustrie, in der überwiegend Verstorbene auftauchen, hat natürlich einen morbiden Beigeschmack.

Beide Techniken haben keine Schnittmengen mit den verwendeten Techniken der Gesichtserkennung, vorausgesetzt es findet ein Austausch des Gesichts durch eine echte Person statt. Der „Puppet-Master" (Marionettenspieler) bietet den besten Ansatzpunkt einer Manipulationsaufdeckung. Für eine Gesichtsüberlagerung braucht es die richtige Mimik, mit der eine bestimmte Person nachgeahmt werden soll. Die Mimik stammt immer vom Imitator und nicht von der zu imitierenden Person. Hier können zusätzlich verfeinerte Ansatzpunkte einer Gesichtserkennung entstehen. Eine derzeitige Mimik-Erkennung bei der Gesichtserkennung wie Lachen, Weinen, Stirnrunzeln etc. reicht für DF-Videos nicht aus. In der Anti-Forensik entstehen derzeit Lösungsansätze, DF zu begegnen (z.B. Open-Source-Toolkit „OpenFace 2.0"). Durch ein Erfassen weicher biometrischer Modelle erfolgt eine Analyse des Gesichtsverhaltens. Ebenso soll ein Erkennen von Fälschungen durch Größenänderungen und Komprimierungen möglich sein.

Komplexität

Wir verlassen nun den Bereich der forensischen Gesichtsanalyse und begeben uns auf die filmwissenschaftliche Ebene. Computergenerierte Filmtechniken haben nicht unbedingt etwas mit Fake zu tun. Deep Fake möchte etwas vorgaukeln, das in Wirklichkeit nicht existiert, z.B. falsche politische Aussagen von prominenten Politikern. Der Einsatz von CGI in der Filmindustrie hat eine ganz andere Motivation. Das Uncanny Valley ist bisher noch nicht gut durchquert worden – also mit computergenerierten Figuren, die dem Menschen so ähnlich wie möglich nachgebildet sind und beim Zuschauer auf Akzeptanz stoßen. Deshalb werden Filme, in denen „menschliche" Roboter oder menschenähnliche Figuren eingesetzt werden, absichtlich übertrieben karikiert dargestellt. Die Ursachen sind vielfältig. U.a. werden Augenblinzeln oder Bewegungen vom Zuschauer als nicht natürlich wahrgenommen. Bewegungen müssten noch viel höher aufgelöst und damit flüssiger gerechnet werden, dementsprechend natürlicher wirken – wir begeben uns da in den Bereich der Supercomputer. Es wird noch eine Weile dauern. **› ei**

Lesestoff

Für 3D Face Reconstruction z.B.:

Automatic 3D face reconstruction from single images or video
▷ **bit.ly/Automatic_reconstruction**

3D Facial model construction and expressions synthesis
▷ **bit.ly/facialmodel**

3D Face Reconstruction from 2D Images
▷ **bit.ly/3D_reconstruction**

OpenFace: an open source facial behavior analysis toolkit
▷ **bit.ly/openface_paper**

Digitale Multimedia Forensik
▷ **bit.ly/thomas_gloe**

Entwicklung einer Gesichtserkennenden Software zur Erfassung und Analyse von Personenströmen
▷ **bit.ly/Michael_Haehnel_1**

Modellbasierte posen- und mimikinvariante Gesichtserkennung
▷ **bit.ly/Michael_Haehnel_2**

Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg
▷ **bit.ly/KI_Buxmann_Schmidt**

Dr. Stefan K. Braun ist seit 2005 als Sachverständiger tätig, er wurde von der Industrie- und Handelskammer 2012 zum „öffentlich bestellten und vereidigten (ö.b.u.v.) Sachverständigen im Bestellungsgebiet Medienproduktion und Mediendesign" ernannt. Ein besonderer Schwerpunkt seiner Tätigkeit liegt in der Audio- und Videoforensik, insbesondere der medienforensischen Erkennung und Rückverfolgung von Raubkopien. Dr. Stefan K. Braun wird von Gerichten, Polizei und Staatsanwaltschaften aus europäischen Ländern beauftragt. 2018 wurde Dr. Braun zum Handelsrichter am Landgericht Frankfurt am Main berufen.
www.medien.expert