

III.2 Umverteilen

Die Ubiquität digitaler Bilder resultiert einerseits aus der medienlogistischen Effizienz ihrer instantanen Distribuierbarkeit. Zugleich ist sie aber Effekt einer anderen Verteilung: der proliferativen Ausbreitung diverser bildtechnologisch begabter, vernetzter Dinge. Die Agency der ›umgebenden‹ Bildinfrastruktur ist über zahlreiche Handlungsträger verteilt, reicht von sensortechnischen Akteuren der *visual capture* bis zu solchen der Datenmobilisierung, -verarbeitung und -visualisierung. Die ›visuelle Kultur‹ des sich ausbreitenden Internets der Dinge folgt Kalkülen, die zwar ebenfalls algorithmisch formalisiert und insofern auf die allgemeine Datafizierung des Bildes angewiesen sind. Gleichwohl wird hier erkennbar weniger logistische Energie in Produktion und Transport bildförmiger Oberflächen investiert.

Auch unter den Bedingungen des »ubiquitous sensing«⁷²⁸ verschwindet das Bild als phänomenale ›Datenaußenseite‹ natürlich nicht vollkommen. Es bleibt eine Option, bestimmte Datenstände anzuzeigen, spezifisch anschaulich und intelligibel zu machen. In aufmerksamkeitsökonomischer wie operativer Hinsicht findet aber eine Verlagerung Richtung Peripherie statt. Dort zirkulieren *image data* als *sensor data*, die mehr oder weniger unmittelbar – also ohne bildhafte Vermittlung – weitergehenden informationstechnischen Verarbeitungsprozessen zugespielt werden. Ganz im Sinne Weisers agiert das verteilte Bild dabei als gedämpftes Signal, das wirksam bleibt, protokollgemäß operative Entscheidungen vermitteln kann, ohne das Volumen des allgemeinen *information overload* weiter zu erhöhen. Als ambiente Signale wirken auf diese Weise distribuierte Bilder strukturgebend, öffnen und schließen Handlungsfelder, treten aber aus sichtbaren Zonen der (herkömmlichen) visuellen Kultur tendenziell zurück.

Während im Rahmen des auf Plattformen und menschliche Nutzer ausgerichteten Internets (›Web 2.0‹) ein übergreifendes Bilddatenverkehrsmodell dominiert – welches die darin zirkulierenden Streams nach den Regelwerken von Computernetzwerkprotokollen überträgt: mittlerweile in ›Real-time‹, potenziell infinit und nach den Vorgaben einer Datenrückflussökonomie –, haben sich die sensorischen Technologien ihrerseits miniaturisiert, ausdifferenziert, verschaltet und bringen durch Datenaustausch- und Datenfusionsprozesse

728 Orit Halpern, Jesse LeCavalier, Nerea Calvillo, Wolfgang Pietsch, »Test-Bed Urbanism«, in: *Public Culture*, 25/2 (2013), S. 273–306, hier: S. 291.

immer kleinteiliger modellierte Umgebungen hervor, die gleichsam schichtweise in Register automatischer Steuerung transferiert werden. Diese benötigt an vielen Stellen bildtechnologisch erfasste Daten, aber nicht unbedingt Bilder im Sinne ihrer Materialisierung als unmittelbar anschauliche Phänomene ikonischer Umsetzung. Ulrik Ekman hat dieses Phänomen als »ambient video« bezeichnet: »Paradoxically, ambient video indicates that ubicomp visibility is best seen without being seen.«⁷²⁹ Bereits Ende der 1980er Jahre äußerte Paul Virilio (vor dem Hintergrund des seiner Meinung nach »virtuell« am Horizont stehenden atomaren Vernichtungskriegs⁷³⁰) die Vermutung, dass »infographische Sehmaschinen«, eine computergestützte, numerische »Visionik« ohne menschlichen Blickkontakt, ohne »videographischen Ausdruck« oder »sichtbaren Zusammenhang« in Zukunft die Aufgabe übernehmen werden, »das umgebende Milieu zu analysieren und automatisch die Bedeutung der Ereignisse zu interpretieren [...]. [...] diese synthetischen Bilder, die von der Maschine für die Maschine hergestellt werden.«⁷³¹

In der Tat: Auf der einen Seite ziehen sich Bilder in »ambientalisierte« Prozesse ihrer automatischen Kalkulation zurück, auf der anderen Seite werden Bildtechnologien zunehmend als »logistical devices of tracking and orientation«,⁷³² als privilegierte sensorische Inputlieferanten beansprucht. Die Informationen, die den solchermaßen instrumentalisierten Bilddaten zu entnehmen sind, lassen sich offenbar von ihrer Zirkulation als visuell umgesetzte »Bilder von etwas« entkoppeln. Für die meisten Sensornetzwerkbilder existieren tendenziell keine Betrachter mehr, sondern nur Transportpläne, Depots und Rechenmaschinen, die sie nach eigenen Gesetzmäßigkeiten auslesen und inzwischen häufig auch als lernalgorithmisches Trainingsmaterial operationalisieren. Es handelt sich also um Bilder, die relevant und folgenreich sind – nämlich operative Mittler für variabel anschließbare Datenverarbeitungsprozesse –, obwohl sie nicht gesehen, sondern lediglich immer weiter verteilt, kalkuliert und durchmustert werden.

Während in Mark Weisers Umgebungsvision die Verteilung zwischen Zentrum und Peripherie noch nachdrücklich auf menschliche Akteure ausgerichtet

729 Ulrik Ekman, »Introduction«, in: ders. (Hg.), *Art and Culture Emerging with Ubiquitous Computing*, Cambridge/MA: MIT Press, 2012, S. 1–59, hier: S. 3.

730 Paul Virilio, *Die Sehmaschine*, Berlin: Merve 1989, S. 152ff. Vgl. dazu auch: Gabriele Schabacher, »Regime der Geschwindigkeit. Paul Virilios Verkehrstheorie«, in: Friedrich Balke, Maria Muhle (Hg.), *Räume und Medien des Regierens*, Paderborn: Fink, 2016, S. 140–167, hier insbesondere S. 154f.

731 Ibid. S. 135ff.

732 Peters, *The Marvelous Clouds*, S. 7.

war,⁷³³ die in fließenden Bewegungen bewusster Aufmerksamkeitsverschiebung entscheiden konnten (und sollten), mit welchen intelligenten, auf einzelne Nutzer individuell »eingestimmten« Informationsflüssen sie situativ verbunden werden wollten, scheint sich nun abzuzeichnen, dass der »periphere Radius« immer größere Zonen enthält, die fundamental intransparent sind. Die »dezenten Signale«, die nach Weiser/Brown durch das gestaltheoretische Vorbild von gleichermaßen verbindenden wie abschirmenden Innenraumfenstern dringen, erreichen menschliche Akteure unter UbiComp-Vorzeichen zwar immer noch. Nur sind die Scheiben mittlerweile im Sinne des *blackboxing* »abgedunkelt«, verunmöglichen also jeden informativen Blick auf Herkunft, Zusammensetzung und Kalkül des Signals einigermaßen grundsätzlich. Je tiefer die Einbettung, je entfernter die Rechenzentren, je komplexer die algorithmische Verarbeitung, Positionierung und Adressierung, desto geringer die Wahrscheinlichkeit, dass die »Schlüsselfunktion der Visualisierung von Information«⁷³⁴ auch nur oberflächlichen Zugang zu den Prozessen der

733 Die bei menschlichen Akteuren verbleibende Kontrolle erläutern Weiser/Brown interessanterweise anhand der »breiteren Kanäle« von Video-Streaming: »Kontinuierlich laufendes Video-Streaming von anderen Orten bedeutet nicht länger Fernsehen oder Videokonferenzen, sondern eher ein Tor zur Aufmerksamkeit. Ein kontinuierlicher Videostream bringt Details in die Peripherie: der Raum ist sauber, vielleicht, weil etwas Wichtiges darin passieren wird; an der Ostküste kamen heute alle spät nach Hause, es muss wohl einen langen Stau oder einen Schneesturm gegeben haben. Multicast teilt mit Video-Konferenzen und dem Fernsehen die wachsende Möglichkeit, sich auf die Wahrnehmung bestimmter Details einzustellen. Verglichen mit einem Telefon oder einem Faxgerät bietet der breitere Kanal volle Multimedia-Möglichkeiten, um Personen besser durch das Kabel hindurch zu »projizieren«. [...] So wie Innenraumfenster ermöglicht es auch Multicast, dass die Kontrolle über die Peripherie beim Individuum verbleibt und nicht ins Environment übergeht.« (Weiser, Brown, »Das kommende Zeitalter der Calm Technology«, S. 68).

734 Ibid., S. 66. Zum (auseinandertretenden) Verhältnis von Vision und Visualisierung schreibt Orit Halpern in *Beautiful Data*: »According to the OED, the term [visualization] is not ancient but rather a modern convention, only appearing in 1883 to depict the formation of mental images of things ›not actually present in sight.‹ Throughout the next few decades, this term expanded to encompass any ›action or processes of making visible.‹ Visualization slowly mutated from the description of human psychological processes to the larger terrain of rendering practices by machines, scientific instrumentation, and numeric measures. Most important, visualization came to define bringing that which is not already present into sight. Visualizations, according to current definition, make new relationships appear and produce new objects and spaces for action and speculation. While the language of vision perseveres, it is important not to assume a direct correlation between vision as a sense and visualization as an object and practice. Married initially to psychology, and now digital computation and algorithmic logic, the substrate and

Formatierung dieser Information vermittelt. Sensorische Verfahren der *visual capture*, die oftmals die entscheidenden Inputdaten für die Auswertung und Steuerung von Umgebungen, aber keine bildhaft-informative Visualisierung mehr erzeugen, sind so gesehen nur ein besonders sinnfälliges Beispiel für komputationsbasierte Automatisierungsprozesse, welche sich in Peripherien zurückziehen, die für Einblicke menschlicher Nutzer allein wegen ihrer mikrozeitlichen Struktur fundamental unerreichbar sind.

Das sich in unterschiedlichen Entwicklungsstufen abzeichnende Internet der Dinge operiert gleichwohl als Netzwerk, das ohne bildsensorische Datenakquisen ebensowenig vorstellbar scheint wie eine zeitgemäße sozialmediale Plattform, deren Nutzungsspektrum darauf verzichten würde, massenhafte, aufwandlose Bildverteilung als zentrales Interface und Praxisoption zu positionieren. Während die mit dem zweiten Komplex assoziierte visuelle Kultur – deren Bildoberflächen dominiert werden von Memes, Selfies, GIFs, von kombinatorischen Mixturen aus neuerfassten und appropriierten Bildanteilen, von bildpraktisch fungiblen Instant-Messaging-Diensten und zahllosen Plattformen des Foto- und Video-Sharings – nur im Hinblick auf die dahinterstehenden Datenkalküle der führenden Medienkonzerne abgeschirmt wird, ist ihre mit immer neuem *image content* bespielte Außenseite keineswegs ›kalmiert‹, sondern durch Formen viralexpanstiver Sichtbarkeit gekennzeichnet. Gespeist aus beschleunigten Distributionspraktiken der Vervielfältigung, Modifizierung, kommunikativen Formatierung von Datenbildern, scheint sich hier eine Tendenz durchzusetzen, die Michel de Certeau Anfang der 1980er Jahre als Symptom der »consumer society« verhandelt und zum Anlass für »opake«, sich der allgemeinen Sicht- und Lesbarkeit entziehende dissidente Spaziergänge genommen hatte: »a cancerous growth of vision, measuring everything by its ability to show or be shown«.⁷³⁵

Unabhängig von kulturkritischen Semantiken ist zweifelsohne zu beobachten, dass die Quantität verteilter Bildsensoren massiv zunimmt, sich in immer dichteren Staffeln durch öffentliche, professionelle, private Räume distribuiert und mittlerweile jenseits des sensorisch multipel begabten

content of this practice has often had little to do with human sense perception or the optic system. Moreover, with the rise of emphasis on haptic interactions and interactivity, visualizations also often take multisensorial modes. Vision cannot be taken, therefore, as an isolated form of perception, but rather must be understood as inseparable from other senses.« (Halpern, *Beautiful Data*, S. 21).

735 Michel de Certeau, *The Practice of Everyday Life*, Berkeley: University of California Press, 1984, XXI.

Smartphones⁷³⁶ in alle möglichen vorgeblich intelligent werdenden Dinge einzieht: von Drohnen über Automobile bis hin zu Werkzeugen, Haushaltsgeräten und Gadgets wie Kleinkinder ausspionierenden Teddybären.⁷³⁷ Die von dieser sensorischen Dissemination ausgehende – stets relational und hervorbringend zu denkende – Umgebungserfassung zieht infrastrukturelle und logistische Systeme, Naturparks, Häuser, Fabrikhallen, Bürotürme, Einkaufszentren, Städte – und nicht zuletzt: sich zusätzlich durch *wearable biometric sensors* »selbstquantifizierende« menschliche Akteure⁷³⁸ – in zunehmend feinmaschigere Netze ihrer environmentalen Verdatung. Auch dank steigender Kapazitäten optischer Auflösung – sowie durch vergleichsweise nachrangige Phänomene wie etwa thermografische,⁷³⁹ panoramatische,⁷⁴⁰ sphärische⁷⁴¹ Expansionsverfahren der *visual capture* – gerät die bildbasierte Erfassung in »Gigapixeln«⁷⁴² immer feinkörniger und lückenloser, wozu, im Sinne einer temporalen Ausdehnung, die mittlerweile bis zu NASAs 4K-Ultra-HD-Streams aus dem Weltall reicht,⁷⁴³ standardisierte Optionen kontinuierlicher Echtzeitübertragung gleichfalls ihren Beitrag leisten. Kurzum: Immer mehr Dinge sind daran beteiligt, immer mehr Weltsegmente ohne Unterbrechungen in Datenbilder zu übersetzen, sie als informatisierte Umgebungen zu konstituieren und operabel werden zu lassen. Sichtbarmachung – »the growth of vision« – zielt auf Steuerbarkeit, Bildtechnologie fungiert als »ausgewachsener Mittler« (Latour), mit der Umgebungen in Daten übersetzt, an Datenströme angeschlossen werden.

736 Vgl. zum Smartphone als sensortechnischer Maschine: Andrejevic, Burdon, »Defining the Sensor Society«; Dourish, Bell, *Divining a Digital Future*, S. 40f. und Greenfield, *Radical Technologies*, S. 30ff.

737 Vgl. amazon.com/zoter.

738 Vgl. Pablo Abend, Mathias Fuchs, Ramón Reichert, Annika Richterich, Karin Wenz (Hg.), *Digital Culture & Society* 2/1 (2016) [Schwerpunkt: »Quantified Selves and Statistical Bodies«].

739 Vgl. Dylan Mulvin, Jonathan Sterne, »Media, Hot and Cold. Introduction: Temperature is a Media Problem«, in: *International Journal of Communication*, 8 (2014), S. 2496–2503 und Lisa Parks, »Drones, Infrared Imagery, and Body Heat«, in: *ibid.*, S. 2518–2521.

740 Vgl. Jay Bolter, Maria Engberg, and Blair MacIntyre, »Media Studies, Mobile Augmented Reality, and Interaction Design«, in: *Interactions*, 20/1 (2013), S. 36–45 und Rothöhler, »Die zwölfte Fläche. Streaming, Mapping, Stitching Places«.

741 Vgl. Elizabeth Woyke, »The 360-Degree Selfie«, in: technologyreview.com, March 2017.

742 Ein Anwendungsbeispiel: »CNN Takes You To The Inauguration of Donald Trump – See the Gigapixel« (in: edition.cnn.com).

743 Vgl. Todd Spangler, »NASA, Amazon to Live-Stream First 4K Ultra HD Video From Space«, in: variety.com, 11.04.2017.

Der Begriff Verteilung deutet in diesem Zusammenhang also nicht zuletzt darauf hin, dass im Kontext diverser Sensornetzwerke neue bildtechnologische Akteure und operative Verknüpfungen entstehen, die in alltägliche Handlungsfelder eingebunden werden – »the ordinary [is] the new frontier of computing«⁷⁴⁴ – und diese gerade durch vermeintlich nichtinvasive Automatisierungsprozesse untergründig zu prägen beginnen, was Kristin Veel im Anschluss an Weisers Terminologie als »calm imaging« bezeichnet hat: »[...] focusing in particular on images that are not addressing our conscious, focused attention, but that operate smoothly in the background. [...] Images emerge and are received calmly as a matter of computational exchanges, but in the process they acquire new potentialities, qualities, and risks [...].«⁷⁴⁵ Im medientechnischen Ambiente sensorischer Umgebungserfassung zirkulieren demnach Bilder, deren ›Sedierung‹ mit einer abweichenden Bewertung ihres Datenstatus zusammenhängt. Als *sensor data* werden Bilder Signale, die tendenziell umweglos instrumentellen Agenden ihrer Berechenbarkeit zufließen oder für spätere Auswertungen in Big-Data-Speichern vorgehalten werden. Um die Signale der *visual capture* produktiv zu machen, bedarf es aus Sicht der verarbeitenden Instanzen keiner bildhaften Phänomenalisierung. Die entsprechenden Kalküle erkennen Muster in rastergrafisch formatierten Werttabellen, in Zahlenmengen, die zwar in vielen Fällen als Bilder encodiert und ausgespielt werden können, dies aber nicht müssen, um effektiv operationalisierbar zu sein.⁷⁴⁶

Während der algorithmische Zugriff auf Bilddaten in Sensornetzwerken auf Mustererkennung abzielt und so gesehen an wahrnehmbaren Bildformen vorbeiführt, lassen sich mit Blick auf neuere Entwicklungen im Feld der sogenannten *computational photography* auch experimentelle visuelle Ausdeutungen dieses Auseinandertretens von Berechnung und Bildhaftigkeit beobachten.⁷⁴⁷ Das Datenbild wird hier nicht ›ruhig‹, sondern auch im Hinblick auf seine visualisierte Außenseite intensiver, invasiver und informativer. So entstehen etwa detailreiche Hochkontrastbilder (*high-dynamic range*) auf Grundlage einer algorithmischen Fusionierung distribuerter bildsensorischer

744 Greenfield, *Everyware*, S. 23.

745 Kristin Veel, »Calm Imaging: The Conquest of Overload and the Conditions of Attention«, in: Ulrik Ekman (Hg.), *Art and Culture Emerging with Ubiquitous Computing*, Cambridge/MA: MIT Press, 2012, S. 119–132, hier: S. 120 und 123f.

746 Vgl. Steven L. Tanimoto, *An Interdisciplinary Introduction to Image Processing. Pixels, Numbers, and Programs*, Cambridge/MA: MIT Press, 2012, S. 3–20.

747 Vgl. Marc Levoy, »Experimental Platforms for Computational Photography«, in: *IEEE Computer Graphics and Applications*, 30/5 (2010), S. 81–87.

Erfassungsvorgänge, die im Fall der avancierten Li6-Kamera des Herstellers Light auf der »folded optics« von bis zu sechzehn sensorischen Akquiseeinheiten basieren. Deren hier nicht mehr sequenziell, sondern simultan erfasste »Einzelbilder« werden zu informationsgesättigten Bildsynthesen »zusammengefaltet«, was einerseits zu signifikant höheren Auflösungen und Kontrastumfängen führt, andererseits aber auch beliebig kombinierbare Schärfebereiche und Optionen nachträglicher Fokusverlagerung ermöglicht.⁷⁴⁸ Konkret materialisierte Bilder enthalten so gesehen zahlreiche latente Schichten an Zusatzinformationen, die als alternative Visualisierungen aus dem gleichen Datenbestand komplementär ableitbar sind.

Ein weitergehender Leistungsschub der Integrationskalküle von Bildbearbeitungssoftware – die in der *imaging pipeline* mittlerweile so nah am Zeitpunkt der Akquise positioniert ist, dass sie echtzeitliche Effekte generiert – wird aktuell vor allem von den plenoptischen Funktionen sogenannter Lichtfeldkameras erwartet.⁷⁴⁹ Deren *smart optic* operiert fundamental logistisch: Erfasst wird nicht einfach Licht, das auf einen Bildsensor trifft, sondern – über Mikrogitterlinsen, die verschiedene Blickwinkel auf eine als Bildfeld gerahmte Szene erzeugen – auch Informationen zu den Richtungen, aus denen die Lichtstrahlen kommen.⁷⁵⁰ Die algorithmische Rückberechnung der jeweiligen Lichttransportstrecken ergibt ein Lichtfeld, dessen raumtiefeninformierte Bilddaten variabel umgesetzt und beispielsweise auch dreidimensional modelliert werden können. Diese unmittelbar im Erfassungsprozess komplex verdatete Lichtlogistik erzeugt somit Bilder, deren konkrete Sichtbarkeit noch

748 Vgl. light.co/technology.

749 Vgl. Ramesh Raskar, »Computational Photography: Epsilon to Coded Photography« [2009], in: web.media.mit.edu.

750 Am Horizont stehen bildtechnologische Lösungen, die visuelle Daten ohne Linsen, ohne optisches System erfassen: »With cameras capturing light differently, a lens isn't necessarily needed anymore. Instead, visual data can be gathered by playing tricks with light, like forcing it through a microscopic grating or diffracting it through a glass sphere. [...] The light passes through a glass sphere instead of a traditional lens, and it is diffracted into a cup of angled sensors. Because the camera knows exactly how light will pass through the sphere, it decodes and stitches the data from each sensor to make a complete image. [...] A group at Rambus Labs is developing lensless image sensors. Instead of the traditional glass, a microscopic grating is placed over the sensor. Light spreads out as it passes through the grating, creating complex patterns on the sensor. The patterns are deliberately cast by the grating, so that the image can be instantly reconstructed by software. Why all the hassle? It's for an entire camera that's less than a millimeter thick. But what we think of as »the picture« would be unrecognizable to the human eye.« (Dave Gershgorin, »Photography Without a Lens? Future of Images May Lie in Data«, in: *The New York Times*, 23.12.2015).

fundamentaler als kontingente informationstechnische Interpretationsleistung prinzipiell umfangreicherer visueller Daten beschreibbar ist: »Computational photography attempts to record a richer visual experience, captures information beyond just a simple set of pixels and makes the recorded scene representation far more machine readable.«⁷⁵¹ Hinter bildhaft umgesetzten Daten stehen hier umfassendere Informationsspeicher, die in Form fundamental anderer Bildausgaben – mit anderen Schärfeverteilungen, anderen Perspektiven usf. – berechenbar sind. Die lichtlogistischen Daten können in phänomenal sehr unterschiedliche Bildformen verpackt werden, welche konstitutiv auf abweichende Umverteilungsoptionen verweisen, die auf der gleichen Datenbasis ebenfalls transportfähig und visuell materialisierbar wären.

Die algorithmisch umstellte *visual capture* generiert hier also wie in Sensornetzwerken nur »intermediate data«,⁷⁵² keine auch nur relativ verbindlichen bildhaften Ausspielungen mehr, wie sie etwa im Verhältnis von *image file* und *image display* gemäß Formaten wie JPEG weiterhin standardisiert sind. Wenn Jennifer Gabrys die vorausgehende Umverteilung sensorischer Perzeptionsleistungen im Anschluss an Katherine Hayles als *distributed sensing* theoretisiert, zielt sie einerseits auf neue Materialitäten und Modalitäten des Wahrnehmens und Erfahrbarwerdens von Umgebungen.⁷⁵³ Menschliche und nichtmenschliche Handlungsinitiative wird in unterschiedlichen Sensornetzwerken flexibel umverteilt, wodurch sich je eigene Umgebungen konstituieren. Von Bedeutung ist hier aber auch, dass mit komputationsfähigen, vernetzten Sensoren nicht nur alternative Datenakquisen und Erfassungsmodi entstehen, sondern zunehmend auch wirkmächtige algorithmische Agenturen instrumenteller Interpretation. Deren *calmness* ist Effekt eines tendenziellen Rückbaus menschlicher Interventionsspielräume:

With these formations of ubiquitous and embedded computing, Hayles suggests that sense is distributed within and automated throughout environments in ways that challenge human sense-making. Sensory processing is not directed through human sensing-subjects primarily but is instead located throughout automated sensing processes. As it is decoupled from human subjects, sensing as a process of making meaning,

751 Ramesh Raskar, Jack Tumblin, Ankit Mohan, Amit Agrawal, Yuanzen Li, »Computational Photography«, in: *The Eurographics Association Eurographics* [diglib.eg.org], 2007, S. 1–16, hier: S. 2.

752 Levoy, »Experimental Platforms for Computational Photography«, S. 81.

753 Vgl. Gabrys, *Program Earth*, S. 63f.

and of generating capacities to make sense and act on information, can occur beyond the realm of human intervention. Indeed, many of the cameras trained on environments are not connected up to human bodies or eyes, but rather work through image analytics to detect patterns and send alerts when significant change is detected. In this case, cameras-as-sensors could be seen to operate in ways that Hayles identifies: as seemingly autonomous agents that are recasting practices of sensing and interpretation.⁷⁵⁴

Menschliche Akteure werden in Sensornetzwerken demnach nicht nur als perzeptive Agenten dezentriert, sondern auch im Hinblick auf die zunehmend verteilte Verarbeitung der sensorisch gesammelten Wahrnehmungsdaten. Aber wie äußert sich diese operative Umverteilung zwischen *sensing* und *sense-making* konkret in bereits außerhalb von Entwicklungslabors antreffbaren Anwendungen?

Vor dem Hintergrund des gewaltigen Anwachsens erreichbarer Bilddatenpeicher scheinen zahlreiche Anbieter vor allem auf Modi interpretativer Vorselektion zu setzen. So kommt im Anwendungsfeld von Sicherheits- und Überwachungstechnologien mittlerweile flächendeckend *video intelligence* zum Einsatz, die die bildsensorisch akquirierten Datenmassen automatischen Analysen unterzieht, um sie pragmatisch handhabbar werden zu lassen. Das typische Leistungsspektrum eines Anbieters wie SeeQuestor zielt hier in erster Linie auf zeitökonomische Effizienzsteigerung. Offeriert wird eine optimierte Durchsuchbarkeit umfangreicher Videodatenbanken: Aus tausenden Stunden Videomaterial, das etwa aus dem polizeilichen Multikamerasystem einer U-Bahn-Station stammt, können Szenen, in denen bestimmte bildinterne Bewegungsmuster enthalten sind, herausgefiltert werden, aber auch eine Galerie der in diesem Bestand erfassten Gesichter. Wird eines davon als »person of interest« markiert, verspricht SeeQuestor eine automatische Extrahierung aller Szenen, in denen die Gesichtserkennungsalgorithmen das gleiche Gesicht detektieren. Ähnliche assistierende Durchmusterungen sollen über »attributes of search« möglich sein, worunter hier die Sortierung der Gesichtsgalerie nach Kriterien wie männlich/weiblich, Brille/keine Brille usf. verstanden wird.⁷⁵⁵

754 Ibid., S. 65.

755 Vgl. seequestor.com. Die interpretativen Prozesse können auch in die Cloud ausgelagert und darüber umverteilt werden, wie das P-React-Projekt verspricht: »Existing surveillance solutions installed are often positioned incorrectly and generate poor image quality unsuitable for evidential purposes. Furthermore, disparate systems coupled with non-existent communication channels, make video exploitation by law enforcement

Die Übergänge zwischen den Anwendungsbereichen sind fließend. So bedient das auf *video analytics* und *smart cameras* spezialisierte kalifornische Unternehmen IntelliVision seine Produktpalette – marktgerecht aufgeteilt in *smart security*, *smart home* und *smart retail* – mit austausch- bzw. erweiterbaren technologischen Lösungen: Bewegungsmelder und Gesichtserkennungs-algorithmen, die besorgte Hausbesitzer über die Aktivitäten von Einbrechern und Haustieren informieren oder bei positiver Identifikation per Datenbank-abgleich das Garagentor öffnen, treffen hier auf videoanalytische Verfahren, die den *consumer flow* in einer Shopping Mall metrisch auswerten und die Bewegungsmuster der Konsumenten in eine *heat map* übersetzen.⁷⁵⁶

agencies nearly impossible. Effective communication and coordination between police, retailers, private security and those operating transport and other infrastructure owners is key. The P-REACT project will design and develop a low cost surveillance platform that will ensure communication between key users with a focus on increasing the ability of on the ground police and security personnel to respond. The solution will encompass intelligent video and audio sensors to detect petty crime incidents, a cloud based monitoring, alert detection and storage platform. Low-cost intelligent sensors (image and audio) will be installed in small business' premises and Transport infrastructure locations. These intelligent sensors will be connected to the cloud-based Video Content Management System (VCMS), where reported incidents will be constantly monitored and responded to. An incident detected by sensors will initiate a work flow including alerting relevant Security Personnel and/or police with the relevant video and intelligence information ensuring the appropriate response. The solution will encourage community participation in the reporting of petty crime and as such will be designed to receive information (images, video) captured by mobile smart devices or unconnected surveillance systems.« (Vgl. p-react.eu).

756 Vgl. intelli-vision.com. In einer Marktstudie erkennt die Unternehmensberatung McKinsey insbesondere im Feld der *retail analytics* die vielversprechendsten Wachstumspotenziale videobasierter IdD-Technologien: »IoT applications usually offer more value when they incorporate video analytics, since the technology allows them to consider a wider range of inputs and make more sophisticated decisions. For instance, some typical IoT applications use beacons that transmit location data each time they connect with a consumer smartphone in a store. While this data can help retailers track the number of visitors, a video-analytics application would provide more detailed demographic information, such as the genders and ages of the shoppers. With their advanced image-processing capabilities, video-analytics applications can consider multiple visual inputs, some of which may be ambiguous and require careful processing. For instance, they can assess the demographics and behaviors of retail customers and turn this information into business insights that assist with product assortment and placement, potentially improving store efficiency, customer conversion, customer loyalty, and other metrics. [...] The software algorithms in video-analytics applications are now capable of gathering and analyzing video footage from multiple sources, thereby generating more detailed insights.

Dass die Optimierung der Identifikation von Objekten und Personen seit einigen Jahren auch bildtechnologische Entwicklungen zu prägen beginnt und dabei »näher« an die sensorischen Erfassungsvorgänge heranrückt, zeigt sich etwa an der »Foveacam« des chinesischen Anbieters Deepglint, die nach der Sehgrube (*Fovea centralis*), dem Bereich schärfsten Sehens bei Säugetieren benannt wurde. Durch ein Verfahren namens »Dynamic Instant Pixel Allocation« kann diese *smart camera* in Bildsegmenten, in denen beispielsweise eine Person oder ein entferntes Autokennzeichen detektiert wird, die Pixeldichte automatisch bis zu 100-fach erhöhen. Die Option eines durch Mustererkennung getriggerten Auflösungszooms verschiebt das Prinzip der Fokusverlagerung in die Übertragungsform eines Livestreams und weitet das sensorisch adressierte, auswertbare Sichtfeld generell aus. Einerseits soll auf diese Weise mehr Raum in den technischen Blick genommen werden, der bei Detektion eines bestimmten Musters instantan und hochaufgelöst komprimiert werden kann. Andererseits wird auch die negative Selektion ereignisloser Videosequenzen automatisiert.⁷⁵⁷ Die Intelligenz derartiger Systeme operiert auf der Grundlage von Monitoring als streamförmiger Dauerübertragung und wird als pragmatisches Gegengewicht zur Proliferation sensorisch erhobener Datenmengen verstanden, soll also durch eine Identifizierung der aus Anwendungssicht instrumentell relevanten Bilddaten Zeit und Speicher ökonomisieren helfen. Der auch für Verfahren des *crowd monitoring* und allgemein der *situational awareness* zentralen Unterscheidung zwischen auffällig/unauffällig, ereignishaft/ ereignislos liegen so gesehen bereits interpretative Skripte in algorithmischer Form zugrunde, die die sensorisch detektierten Daten für weitere Bedeutungs-zuweisungen vorselektieren.

Die informationstechnische Auswertung derartig einströmender Bilder erfolgt, um die meisten von ihnen distributiv ausschließen zu können, sie gerade nicht weiterverteilen zu müssen. Ihr Status als ungesehene Bilder, als sensorische Outtakes, ist in vielen Fällen gleichwohl temporär, sofern sie zumindest mittelfristig in (fast immer zentralisierten) Speichern bleiben, aus denen zu späteren Zeitpunkten andere Sortierungen generiert werden können. Solange die Reste nicht gelöscht werden, bleiben sie zwischengelagert, im Ambiente, bilden sensorische, in die Umgebungsgeschichte führende Datenarchive, die

For example, surveillance applications can identify people based on physical characteristics from video feeds collected at multiple locations at different times. Similarly, retail applications can aggregate data from multiple video feeds to determine the shopping patterns characteristic of different demographic groups.« (Vasanth Ganesan, Yubing Ji, Mark Patel, »Video meets the Internet of Things«, in: mckinsey.com).

757 Vgl. deepglint.com.

auch in entfernteren Zukünften korrelativ und prädiktiv kalkulierbar sein sollen:

[...] ›sensor society‹ thus refers not just to the proliferation of automated sensing devices across the landscape, but also to the associated logics that are characteristic of automated, mechanized sensing: always-on information capture, the associated avalanche of data, and the consequent tendency toward automated information processing and response. It also refers to the changing ways of thinking about and using information with which automated sensing is becoming associated: in particular the Janus-faced use of data for both prediction and information retrieval – that is, the related goals of modelling the future and reconstructing the past.⁷⁵⁸

Während Anwendungen wie Googles Cloud Video Intelligence mit Lernalgorithmen daran arbeiten, die Bilddatenspeicher für Verfahren der ›Volltextsuche‹ zu öffnen, Bildinhalte (›label detection‹) oder formale Eigenschaften (›shot change detection‹) automatisch zu erkennen und zu annotieren,⁷⁵⁹ sind Sensornetzwerke zunehmend auf *real-time analytics* angewiesen. Deren pragmatische Umsetzung strukturiert immer mehr bildtechnologischen Alltag: Tritt ein neuer Postbote an den Briefkasten, wird ein ›unfamiliar face‹ rapportiert, was allerdings auch passieren kann, wenn das Aufnahmegerät etwas ratlos auf sein eigenes Spiegelbild trifft.⁷⁶⁰ Ergibt ein Datenbankabgleich, dass das richtige,

758 Andrejevic, Burdon, »Defining the Sensor Society«, S. 13f.

759 Vgl. cloud.google.com/video-intelligence und Fei-Fei Li, »Announcing Google Cloud Video Intelligence API, and more Cloud Machine Learning updates«, in: cloud.google.com/blog, 08.03.2017.

760 »To test the Nest Cam IQ, I set it up in the Financial Time's San Francisco bureau a few weeks ago. Every morning I got alerts telling me when my colleagues arrived in the office and when our regular cleaner came in after we had all left for the night. If someone arrived that the camera had not seen before, the Nest app warned me about an ›unfamiliar face‹. If I knew the person, I could enter their name; next time Tom or Hannah appeared, it would tell me so by name. In a small household, this may not be particularly useful. But for a large family it could bring peace of mind that the kids all got home safely, for instance, or warn when a stranger appears unexpectedly. If you do not want to get a push notification for every single person's comings and goings, the app collects moments of activity into a timeline for review at a later date. Nest's face-recognition software was pretty accurate, if not 100 per cent reliable. It did not eliminate false positives altogether, although it reduced them to a handful a week. When placed in one location, propped up on a tripod, I realised that it was mistaking its own reflection in a meeting room's glass wall for a human head and shoulders. [...] Other new features that mark the Nest Cam

per Datenbild autorisierte Gesicht vor der Haustür steht, wird diese geöffnet, ohne dass sich ein Betrachter mit der kontinuierlichen Bildproduktion des Home-Security-Systems beschäftigen müsste. Erkennt eine *smart camera*, dass ein bestimmter Koffer für längere Zeit unbeaufsichtigt auf einem Bahngleis steht, wird ohne menschliche Beurteilung des Streambildes sicherheitshalber Alarm ausgelöst: »These images are not fixed representations of environments, but rather are temporal markers of visual sensing data within larger data sets [...]«. ⁷⁶¹ Die intelligente Auswertung eines Videofeeds soll Personen, Objekte, Situationen echtzeitlich – in bestimmten Überwachungssystemen eigentlich »antizipatorisch« ⁷⁶² – detektieren, also in Bezug auf die einströmenden Bild-daten laufend Unterscheidungen vornehmen, die definieren, was als relevantes Signal gilt, im Arbeitsspeicher und anzeigbar bleiben sollte, und was als Noise zu bewerten ist, gelöscht oder in entferntere, weniger erreichbare Speicher ausgelagert werden kann.

Die dabei entstehenden, von nichtmenschlichen Akteuren quasi-autonom verwalteten Bildkreisläufe – und die Effekte der darüber operationalisierten Sortierleistungen – hat Kristin Veel am Beispiel von Smartgate beschrieben, einem in australischen Flughäfen installierten Sicherheitssystem, das mit Gesichtserkennungssoftware arbeitet:

[Smartgate] takes an image of a traveller's face and matches it with a digitized image stored in the traveller's ePassport. If the match is successful, the traveller is cleared through customs and immigration control. The

IQ out over the original device include the ability to zoom in when you are viewing live footage, thanks to that 4K sensor. Its Supersight software means that the video pans and zooms to follow a person around the room, even though the camera itself does not physically move.« (Tim Bradshaw, »Review: Nest's Sharp Upgrade to Home Security Camera«, in: *The Financial Times*, 14.07.2017).

761 Gabrys, *Program Earth*, S. 69.

762 »Crucial to these emerging surveillance systems is a radically new politics of anticipatory seeing. For the overarching feature of the new, militarized, surveillance push, whether its ›targets‹ are located in Manhattan and Baghdad, London and Fallujah, is an attempt to build systems of technological vision in which computer code itself is, along with data-bases of real or imaged ›targets‹, delegated with the agency of tracking and identifying ›abnormal‹ ›targets‹ from the background ›normality‹ of a homeland or war-zone city. Crucial here is the adaptation of the commercial practices of ›data mining‹ or ›predictive analytics‹ where algorithms are developed to look for patterns in the swathes of captured data, identify or profile behaviours or characteristics deemed to be ›unusual‹ or ›abnormal‹, and search for ›target‹ people, transactions or flows deemed to have such characteristics.« (Crang, Graham, »Sentient City«, S. 803).

technique obviously allows for fast processing of the increasing number of travellers in a globalized world. However, it is also likely to have a significant impact on travelling patterns and on what has been termed ›software-sorted societies‹. [...] In the case of software sorting, we are thus seeing an example of calm surveillance technologies, which are undeniably countering the risk of overload inherent in border controls, but which have a huge impact on people's lives and the structure of society. [...] The images involved in such digital identity systems are most often not taken by and not received by humans.⁷⁶³

Die *calmness* der Sensornetzwerkbilder resultiert also aus den erweiterten Möglichkeiten, sie ungesehen auszuwerten, zu verwerfen oder entlang einer operativen Kette zirkulieren zu lassen und schließlich in beliebig anschließbare Entscheidungsautomatismen eines »computational exchange« (Veel) zu übersetzen. Agency wird in diesem Szenario tendenziell zugunsten nicht-menschlicher Bildakteure umverteilt, die nicht nur zunehmend autonome Operationen der *visual capture* umsetzen – und insofern, wie Mark B.N. Hansen verschiedentlich ausgeführt hat, ein sensortechnisch konstituiertes »umweltliches Empfindungsvermögen« ausbilden⁷⁶⁴ –, sondern auch beginnen, aus den visuellen Umgebungsdaten ihre eigenen bzw. vorprogrammierte Schlüsse zu ziehen. Aktive menschliche Handlungsbeiträge werden hierbei weder für die Datenerhebung, noch für deren Verarbeitung benötigt.

Für Hansen sind die daran beteiligten bildsensorischen Verfahren nicht mehr über die Untersuchung einzelner Bildobjekte beschreibbar, weil diese sich – darin der Trajektorie videografischer Diskretisierung folgend – in unwahrnehmbare mikrotemporale Prozesse zerlegen: »What appears beneath the shards of this exploded image is the dynamic and constructive, microtemporal process of imaging or modulation that directly engages and synthesizes (contracts-dilates) time matter itself, without the assistance of the form of the image and prior to the operation of macroconscious perception.«⁷⁶⁵ Bewusste Wahrnehmungsleistungen, die sich auf relativ stabile sinnliche Einheiten wie etwa *image content* beziehen, werden demnach durch heterogene sensorische

763 Veel, »Calm Imaging«, S. 122.

764 Mark B.N. Hansen, »Medien des 21. Jahrhunderts, technisches Empfinden und unsere originäre Umweltbedingung«, in: Erich Hörl (Hg.), *Die technologische Bedingung. Beiträge zur Beschreibung der technischen Welt*, Frankfurt/Main: Suhrkamp, 2011, S. 365–409.

765 Mark B.N. Hansen, »Ubiquitous Sensation: Toward an Atmospheric, Collective, and Microtemporal Model of Media«, in: Ulrik Ekman (Hg.), *Art and Culture Emerging with Ubiquitous Computing*, Cambridge/MA: MIT Press, 2012, S. 63–88, hier: S. 67.

Informationsflows ersetzt, die zwar auch »Sinnlichkeit« distribuieren, diese aber nicht mehr »makroskopisch« formatieren und hochskalieren:

For the first time in our history and (very likely) in the history of the universe, we find our long-standing and up to now well-nigh unquestionable privilege as the world's most complex sensing agents challenged, if not overthrown, by the massively replicable and ubiquitously propagating technical capacity for sensing introduced by our smart devices and technologies. Taken individually, these devices and technologies are of course vastly less complex – by many orders of magnitude – than human body-minds; cumulatively, however, and with their capacity to operate almost entirely without interruption and across a vast range of scales, they have already begun to dwarf us in their ability to gather and generate sensory data.⁷⁶⁶

Mindestens der Hinweis auf eine qualitativ neue Reichweite umgebender sensorischer Datenströme, die nicht mehr als sinnlich anschauliche, lesbare Formen ausgegeben werden müssen, um auf menschliche Akteure einzuwirken,⁷⁶⁷ scheint mit Blick auf automatisierte Prozesse in Sensornetzwerken, die Bilddaten analytisch durchmustern und gegebenenfalls operative Signale aussenden, durchaus plausibel. Die diagnostizierte Deprivilegierung menschlicher Agency gerät in derartigen Befunden zum Resultat einer medientechnischen Expansion und Optimierung: Einerseits geht die sensortechnische Durchdringung der Umgebung mit immer weiter ausdifferenzierten Modi ihrer

766 Hansen, *Feed-Forward*, S. 161. Ulrik Ekman bemerkt zu diesem Punkt: »[Mark Hansen] delineates the sensory impact of ubicomp as a question of an expansion of the visible to include the imperceptible: a new correlation of media and sensation that occurs beneath the temporal frame definitive of both image and conscious perception. Hansen approaches the invisibility of ubicomp as the invocation of a microsensational address by the image in its technical expansion to include a periphery of imperceptible but objectively sensory informational flows.« (Ekman, »Introduction«, S. 25f.).

767 Jennifer Gabrys versucht, diese Einwirkungen mit Foucaults spätem Hinweis auf eine »umgebungsbezogene Regierungsmacht« zu analysieren (vgl. Gabrys, *Program Earth*, S. 190). In *Die Geburt der Biopolitik*, in der für diesen Diskurs einschlägigen Fußnote zu einer der letzten Vorlesungen Foucaults, heißt es: »[...] wir [haben] in diesem Horizont das Bild, die Idee oder das politische Thema einer Gesellschaft, [...] in der es keine Einflussnahme auf die Spieler des Spiels, sondern auf die Spielregeln geben würde und in der es schließlich eine Intervention gäbe, die die Individuen nicht innerlich unterwerfen würde, sondern sich auf ihre Umwelt bezöge.« (Michel Foucault, *Geschichte der Gouvernementalität II: Die Geburt der Biopolitik, Vorlesung am Collège de France 1978/79*, Frankfurt/Main: Suhrkamp, 2004, S. 359).

informatisierten Steuerbarkeit einher. Andererseits führt die Automatisierung, die mikrozeitliche Taktung und die in ambienten Hintergründen verborgene Logistik der entsprechenden Prozesse – die die ubiquitär erhobenen Umgebungs- und Verhaltensdaten transportieren, kalkulieren, bewältigen soll – zu einer Verringerung wahrnehmbarer, pragmatisch erreichbarer Interventionspielräume. Die skizzierte Umverteilung des Bildes ergibt sich unter diesen Voraussetzungen also vor allem im Hinblick auf die zunehmende Bedeutung von Bildakteuren, die Umgebungsdaten sensorisch erfassen und algorithmisch prozessieren, ohne dass der visuelle Input als bildhafte Form stabilisiert und angezeigt werden müsste.

Diese tendenzielle ›Bildlosigkeit‹ bildtechnologischer Einsätze – wie gezeigt: verstanden als Effekte einer Umverteilung, die phänomenale Bildmaterialisierungen zugunsten automatisch weitertransportierter Bilddaten zurückstellt⁷⁶⁸ – manifestiert sich im Anwendungskontext der sogenannten Augmented Reality (AR) in gewisser Hinsicht ambivalenter. Dort fungieren Bildsensoren zwar ebenfalls als privilegierte Inputgeräte und Mittler, erzeugen aus Nutzerperspektive aber zugleich informationsgesättigte Interfaces, die nach wie vor bildhaft formatiert sind. So heißt es in einem einschlägigen anwendungsorientierten Handbuch: »Virtuelle Objekte werden in reale, durch Kameras bereitgestellte Szenen in Echtzeit so eingefügt, dass sie räumlich korrekt positioniert sind und das reale Bild ergänzen. Die digitale Information verschmilzt mit der Umwelt des Benutzers; dies ermöglicht, dass der Nutzer die aktuell wichtigen

768 Abgesehen von den eingangs bereits zitierten Positionen von Claus Pias und Birgit Schneider lässt sich auch mit Wolfgang Hagen in Bezug auf den immanenten »Ikonoklasmus« (Gottfried Boehm) digitaler Bilder noch grundsätzlicher – nämlich im epistemologischen Sinn – von deren »spielerischer« Bildlosigkeit sprechen: »Nicht also wegen der Technik als Hardware an sich, sondern wegen der Epistemologie, die in der Hardware der Chips beschlossen liegt, mündet meine These, dass es ein ›digitales Bild‹ nicht gibt, in dem Satz: Im Digitalen gibt es immer nur die Chance auf ein Bild. Bilderzeugung und Bilddarstellung auf elektronischem Wege ist nichts anderes als ein nicht-destruktives Chancenspiel. Das resultiert nicht einmal ursächlich aus dem Vorgang der Digitalisierung. Die Digitalisierung rechnet zwar die Elektronenladungen in Bits und Bytes um, aber setzt erst ein, wenn der Chip nach den Chancen des quantenmechanisch Berechenbaren belichtet wurde. Die Digitalisierung freilich führt das isomorphe Spiel der Chancen fort, und zwar als Spiel, aus dem es kein Entrinnen mehr gibt, weil es prinzipiell und epistemologisch der Entropie nicht unterliegt. Bis ein digitales Bild entsteht, müssen auf jeder Ebene Entscheidungen getroffen werden. Im Chip sind alle Elektronen unbildlich. Die aus ihnen gewonnenen Bits kann man nur zählen.« (Wolfgang Hagen, »Es gibt kein ›digitales Bild‹ – Eine medienepistemologische Anmerkung«, in: *Archiv für Mediengeschichte*, 2 (2002), S. 103–112, hier: S. 110).

Informationen direkt an dem Ort erhält und sieht, an dem er sie benötigt. [...] Die Einblendung erfolgt kontextsensitiv, d.h. passend und abgeleitet vom betrachteten Objekt. So wird das reale Sichtfeld beispielsweise eines Kunden durch eingeblendete Produkthinweise oder -darstellungen um für ihn wichtige Informationen erweitert.«⁷⁶⁹

In Abgrenzung zur Virtual Reality (VR)⁷⁷⁰ geht es bei AR um eine Integrationsleistung, die computergenerierte Informationen oder Zusatzobjekte (in letzterem Fall wird ein »dreidimensionaler Bezug zwischen realen und virtuellen Objekten generiert«⁷⁷¹) nicht in eine synthetische Umgebung, sondern in eine »Realität« transportiert, die insofern als »reales Bild« zu verstehen ist, als sie live auf einem bildsensorisch erfassten Sichtfeld erscheint.⁷⁷² Die räumlichen Grenzen eines derartigen Livefeeds entsprechen somit den Randverläufen einer kontinuierlich zwischengespeicherten, videografisch erzeugten Bildfläche, welche, etwa bei Anwendungen auf der Basis mobiler Handhelds,⁷⁷³ hinsichtlich der Kadrate ihrerseits beweglich, vorläufig ist.

769 Anett Mehler-Bicher, Michael Reiß, *Augmented Reality: Theorie und Praxis*, Berlin: De Gruyter, 2011, S. 2.

770 Paul Milgram et al. haben AR als »mixed reality«-Kontinuum definiert, wobei die außerhalb dieser variablen Mischverhältnisse liegenden Pole als »real environment« und »virtual environment« bezeichnet werden: Vgl. Paul Milgram, Haruo Takemura, Akira Utsumi, Fumio Kishino, »Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum«, in: *Telemanipulator and Telepresence Technologies*, Vol. 2351 (1994), S. 282–292.

771 Mehler-Bicher, Reiß, *Augmented Reality*, S. 6.

772 AR ist meist, muss aber nicht auf Sehleistungen basiert sein: »[...] we do not consider AR [...] to be limited to the sense of sight. AR can potentially apply to all senses, augmenting smell, touch and hearing as well. AR can also be used to augment or substitute users' missing senses by sensory substitution such as augmenting the sight of blind users or users with poor vision by the use of audio cues, or augmenting hearing for deaf users by the use of visual cues.« (Julie Carmigniani, Borko Fuhr, »Augmented Reality: An Overview«, in: Borko Fuhr (Hg.), *Handbook of Augmented Reality*, New York: Springer, 2011, S. 3–46, hier: S. 3).

773 Greenfield versteht AR als verteilte »interface technique« und beschreibt das dahinterstehende Prinzip folgendermaßen: »Bundle a camera, accelerometer/gyroscope, and display screen in a single networked handset, and what you have in your hands is something that can sustain at least a rudimentary augmentive overlay. Add GPS functionality and a high-resolution three-dimensional model of the world – either maintained onboard the device, or resident in the cloud – and a viewer can be offered location-specific information, registered with and mapped onto the surrounding environment. In essence, phone-based AR treats the handset like the transparent pane of a cockpit head-up display: you hold it before you, its forward-facing camera captures the field of view, and an overlay of information is applied on top of it. Turn, and the onscreen view turns with you, tracked (after

Übertragung und Prozessierung erfolgen instantan, in unwahrnehmbaren mikrozeitlichen Taktungen medientechnisch vermittelter ›Real-time‹, stehen aber auch hier mit medienlogistischen Speicherprozessen in Verbindung: einerseits mit Datenbanken, in denen die eingeblendeten Informationen – wie etwa virtuelle Zusatzobjekte – als Datensätze hinterlegt sind, andererseits mit jenen Datenrücksendungen, die insbesondere in kommerziellen Anwendungen integraler Bestandteil der Lieferprotokolle sind. Auch für das vermeintlich ganz in seiner Liveness und Transparenz aufgehende AR-Interfacebild gilt folglich, was John Durham Peters als »a boom in data« bezeichnet hat: »In an online world every act leaves a trace, a record of some sort, and such documentation provides potent data to those who can access and read it.«⁷⁷⁴

Die auf dieser navigierbaren Datenfläche bildhaft umgesetzte ›Realität‹ wird zwar durch virtuelle Objekte (derzeit besonders beliebt: Möbel und Maßbänder⁷⁷⁵) ergänzt bzw. überblendet, aber nicht, wie in VR-Applikationen, komplett abgeschaltet. In ortsbasierten AR-Games – von ARQuake (2000) bis zu Pokémon Go (2016) –, aber auch in medizintechnischen Assistenzanwendungen, die ›Röntgenblicke‹ simulieren,⁷⁷⁶ oder bei AR-Einsätzen zur Wartung komplexer Industrieanlagen erfolgt die visuelle Ausgabe (*rendering*) mit dem Ziel, Interaktionen in Echtzeit zu ermöglichen, was in vielen Fällen das Erkennen und Tracking von Objekten im umgebenden Realraum voraussetzt. Bevor eine Information oder ein computergrafisches Objekt als »Realitätserweiterung« in das Sichtfeld eingerechnet und als augmentiertes Livebild gerendert werden kann, müssen in diesem trackbare Muster erkannt worden sein (als Hilfsmittel für minimalinvasive Chirurgie möglichst millimetergenau): »Computer vision renders 3D virtual objects from the same viewpoint from which the images of the real scene are being taken by tracking cameras. [...] fiducial markers, optical images, or interest points are detected in the camera images. Tracking can make use of feature detection, edge detection, or other image processing methods to interpret the camera images.«⁷⁷⁷ Bildsensoren operieren hier als lokalisierende *pointing devices*, durch deren (informativ erweiterte) Bildprodukte gleichsam ›hindurchgesehen‹⁷⁷⁸ werden soll.

a momentary stutter) by the grid of overlaid graphics. And those graphics can provide anything the network itself offers.« (Greenfield, *Radical Technologies*, S. 159f.).

774 Peters, *The Marvelous Clouds*, S. 7.

775 Vgl. Jason Tanz, »Apple Bets the Future of Augmented Reality Will Be on Your Phone«, in: wired.com, 08.07.2017.

776 Mehler-Bicher, Reiß, *Augmented Reality*, S. 16.

777 Carmigniani, Fuhr, »Augmented Reality«, S. 6.

778 Carmigniani/Fuhr sprechen hier von »video-see-through technique« (ibid., S. 10).

Das algorithmisch erzeugte und als Interface beanspruchte AR-Bild tritt gerade nicht hinter automatisierten Kalkülen zurück, sondern wird zu einem offenen, standortbewussten, permanent bildhaft visualisierten Informationsspeicher, dem unter Echtzeitbedingungen Muster extrahiert und hinzugefügt werden können: »[...] the once self-contained image, at most supplemented or anchored by textual information can now be overlaid or augmented with visual, textual or aural information individually tailored to the viewer's exact location and specification.«⁷⁷⁹ Bildsensoren verwandeln sich dabei in gewisser Weise in räumlich distribuierte Suchmaschinen, die in der ›realen‹ Welt nach datafizierten Kontaktpunkten für die Ausspielung computergenerierter Informationsflüsse fahnden, was in jüngster Vergangenheit – etwa mit Blick auf echtzeitliche Bilderkennungsverfahren wie Googles noch nicht marktreife »visual search app« Lens, die in den virtuellen Assistenten des Anbieters integriert werden soll – bereits Anlass zur Vermutung gab, dass nicht allein Sprachassistenzsysteme, sondern zunehmend auch Kameras dabei seien, das Keyboard zu ersetzen: »Point your camera at a tree, and it'll tell you the variety. Snap a pic of the new restaurant on your block, and it'll pull up the menu and hours, even help you book a reservation. Perhaps the single most effective demonstration of the technology was also its dullest – focus the lens on a router's SKU and password, and Google's image recognition will scan the information, pass it along to your Android phone, and automatically log you into the network. This simplicity is a big deal. No longer does finding information require typing into a search box. Suddenly the world, in all its complexity, can be understood just by aiming your camera at something.«⁷⁸⁰ Auf W.J.T. Mitchells eingangs zitierte Frage – »What Do Pictures Want«? – antwortet Googles Lens mit dem automatisch eingelesenen, Konnektivität herstellenden Bild eines Routerpasswortes, also einigermaßen unzweideutig: Digitale Bilder wollen ins Netz, anschlussfähig, verteilbar werden.

In seiner augmentierten Variante ist das verteilte Bild gleichwohl nicht ›calm‹ – wie, zumindest der Tendenz nach, in Sensornetzwerken –, sondern eine visuell ›angereicherte‹ Datendarstellung, die auf der Oberfläche nach wie vor tradierte fotografische Konventionen übermittelt, zugleich aber als Schnittstelle für informationstechnische Tiefenschichten operabel wird. Die instantan prozessierten Datagramme der *visual capture* fließen algorithmisch gesteuert Rechenzentren zu, die die eingehenden Datenpakete (zwischen) speichern und mit geeigneten Speicherbeständen – ohne zeitlichen Verzug

779 Kember, »Ambient Intelligent Photography«, S. 57.

780 Elizabeth Stinson Gear, »Your Camera Wants to Kill the Keyboard«, in: wired.com, 22.05.2017.

und passgenau – beantworten, was zu weiteren Anfragen, neuen Kalkulationen, Horizonten fortgesetzter Distribution führt. Bildtechnologie operiert innerhalb der Infrastrukturen und Protokolle dieses environmentalen Datenverkehrsmodells als Mittler, der »machinic access to the data of sensibility«⁷⁸¹ bereitstellt und verarbeitet, um als *field of vision* sensorisch erfasste Umgebungen durch dynamische Filter fortlaufender Mustererkennung zu ziehen. Das verteilte Bild folgt somit auch hier einem logistischen Auftrag: Es liegt als Prozess im Sucher, zirkuliert in Sensornetzwerken, deponiert verdatete Perzepte und stellt informative Transportanschlüsse her.

⁷⁸¹ Hansen, *Feed-Forward*, S. 53.