Das digitale Bild und Neue Visuelle Technologien

Inhaltsverzeichnis

Einleitung	2
1. Teil: das digitale Bild	2
Die digitale Revolution	2
Das digitale Bild ist berechnet...	3
Spezifische Zeitlichkeit und Räumlichkeit des digitalen Bildes	3
Numerische Bilder	5
Visuelle Technologien	6
Kritik des Visuellen	7
Technik ist stabilisierte Gesellschaft	7
2. Teil: Neue Visuelle Technologien	10
Algorithmic Turn	11
Photosynth	12
Augmented Reality	13
Zwei Beispiele	15
Mitsubishi Electric	15
Google Glass	16
Ausblick und Schluss	17
Literatur	18

Einleitung

In dieser Arbeit soll es sich um *Digitalisierung* drehen, und zwar im Besonderen um – und mit Beschränkung auf – die *Digitalisierung des Visuellen*. Ich möchte in einem ersten Schritt einige Aspekte aufzählen und beschreiben, die mir in diesem Bezug für eine Soziologie des Visuellen als wichtig erscheinen; später, in einem zweiten Schritt, werde ich *Neue Visuelle Technologien* in den Blick nehmen, und zwar mit einer besonderen Obacht auf solche, die im Kern von *Algorithmen* bewegt und strukturiert werden. Gestützt auf Uricchio (2011) möchte ich argumentieren, dass bei der Interpretation von Bildern bzw. – allgemeiner – des Visuellen – eine genaue Beschreibung jener softwareimmanenten, automatisierten Verfahrensweisen als *Movens* der auf die Neuen Visuellen Technologien bezogenen Wahrnehmungs- und Handlungsweisen für ein soziologisches Verständnis gegenwärtiger – und zukünftiger – Alltagspraktiken notwendig ist.

Ich werde dazu – zunächst mit Hilfe von Urrichio – einige Beispiele für Algorithmus-basierte visuelle Technologien beschreiben; und schließlich eine Annäherung an die aktive Rolle – mit der Actor-Network-Theory gesprochen: an die *agency* solcher automatisierter Software-Routinen im Herzen der neuen visuellen Technologien versuchen, in einer Welt mit rasend zunehmender mediatisierter Visualität.

1. Teil: Das digitale Bild

Die digitale Revolution

Mitchell zählt in einem Vortrag¹ drei Revolutionen auf, die die Bilderwelt bzw. eine Kultur des Visuellen in ihren Grundfesten erschüttert und transformiert hätten: Da wäre erstens den mit Benjamins '*Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*' beschriebenen Auraverlust von Bildern durch ihre eben technische Reproduzierbarkeit; zweitens die *digitale Revolution*, die seit einigen Jahrzehnten andauert und – ebenso wie erstere, 'analoge', Revolution, auf allen Gebieten der menschlichen Produktion, Konsumtion und Kommunikation für neue Strukturen gesorgt hat und sorgt; und drittens die von Mitchell sogenannte

¹Vortrag von W.J.T. Mitchell im Haus der Kunst München, am 9.6.2012; abrufbar auf: <http://www.youtube.com/watch?v=c-XdkcBkPY&list=PLaQauau23m3PV69c6ax1Mv3NxbPjOYkqD&index=16>, 30.04.2013, 10.00

biokybernetische Revolution, die es neuerdings erlaube, Kopien bzw. Klone von Lebewesen zu erzeugen.

Ich möchte mich hier ausschließlich über Zweiteres ausbreiten. Da ich denke, dass es gerade auf diesem Gebiet in den letzten Jahren – Stichwort automatisierte visuelle Erkennung und Augmented Reality - qualitativ und auch quantitativ ungeheuer transformative Entwicklungen gegeben hat und dies sich in der Gegenwart als recht relevant erweist: insbesondere auf dem Gebiet des normalisierten Alltagshandelns.

Das digitale Bild ist berechnet...

Digitale Bilder basieren auf Binärcode – Nullen und Einsen. Licht wird bei ihrer Produktion in Code übersetzt – und muss bei jeder Darstellung für das menschliche Auge wiederum über mehrere Stationen in Licht rückübersetzt werden. Es ist also immer in der einen oder anderen Weise ein Rechner erforderlich, der das in Binärcode gespeicherte Bildformat in in bestimmter Weise arrangierte Bildpunkte umrechnen kann.

Und nun – in der Gegenwart - ist die potentielle Rechenleistung so groß, dass mensch bzw. maschine inzwischen während dieses Berechnungsvorgangs diesen so schnell verändern kann, sodass – in ‚Echtzeit‘ – z.B. ein Amalgam aus materiellen Repräsentationen und virtuellen Kreationen zustande kommt, bei dem immer weniger klar wird, welche der Bildelemente denn nun empirische Referenz aufweisen, und welche denn nun vollkommen errechnet sein müssen.

Dies wird mich im Kapitel zur ‘Augmented Reality’ später noch beschäftigen.

Spezifische Zeitlichkeit und Räumlichkeit des Digitalen Bildes

Bilder sind ubiquitär:

heißt nicht nur, dass Bilder überall hinkommen kommen, sondern dass sie *schon da sind*: Z.B. Google Earth und – Maps für die Vogelperspektive; unzählige Überwachungskameras für die Straßenperspektive, Bilder-teil-programme wie Instagram für die (‘freiwillige’ Aufgabe der) Intimsphäre, und das eigene Gesicht; vom Fernsehen und Video gar nicht zu sprechen;

bildgebende Verfahren in der Medizin; und schließlich – aktuell *Google Glass*–TrägerInnen als BildproduzentInnen, die ihre eigene visuelle Perspektive potentiell gleichzeitig an unzählige andere – Nicht-Anwesende – senden können. Und das in...

...*Echtzeit*:

Untrennbar verwoben mit dem Phänomen der Ubiquität ist die Zeitlichkeit der *Nullzeit* “Die neuesten Technologien lassen den Raum in seiner Ausdehnung und Dauer verschwinden. Sie reduzieren die Welt auf ein Nichts...” (Virilio 1995) sagt Virilio. Ob dies von so großen, welthistorischem Nachteil ist, wie Virilio behauptet, ist im Zusammenhang der neuen visuellen Technologien nach wie vor in Diskussion.

Ubiquität und Echtzeit bzw. Nullzeit sind möglich durch die zunehmenden Verbindungen und Verbindungsgeschwindigkeiten – der Bandbreite – des Internets und natürlich prinzipiell durch die unendliche Möglichkeit des Kopierens von Binärcode – und dem sofortigen Versenden desselben. Wenn Lady Gaga ein Bild von sich auf Twitter postet, können dies – aktuell – 36 Millionen ‘Follower’ - im selben Augenblick sehen, wo dieses Bild gemacht und online gestellt ist. Und zwar egal, ob sie auf der Internationalen Raumstation, in der Antarktis, in New York oder Wien surfen.

Doch diese Bilder sind im Vergleich zu ‘materiellen’, analogen Bildern auch sehr viel *flüchtiger*: man betrachtet das Bild, und lässt es, meist Sekunden später, unsichtbar werden, im Speicher verschwinden, sodass es nur mehr virtuell als Code irgendwo in irgendeinem Speicher auf der Welt vorhanden ist; oder eher noch wird das Bild gelöscht, oder im temporären Speicher vergessen und irgendwann automatisch vom Computer gelöscht. Glück muss es haben, dass es schon so weit kommt: wenn wir digital Fotos machen, verwerfen wir schon gleich einmal ein paar, bevor einem erlaubt wird, dass mit ihm weitergerechnet werden kann.

Zur eingehenden Verdeutlichung der Flüchtigkeit von digitalen Bildern könnte ein Vergleich der Fotonotizfunktion von Apps wie *Evernote* mit der vorherigen Photographierpraxis mit analogen Apparaten dienen. Mit letzteren machte man ein Foto – und dann wars das auch schon erstmal. Weder konnte man es sofort auf einem Bildschirm sehen, noch konnte man es löschen, wenn es nicht beliebte, es sei denn, man machte den Apparat im Sonnenlicht auf,

und ruinierte dabei den ganzen Film auf immer und ewig. Fertige Fotos kamen dann in der Regel einige Tage danach, nach der Bearbeitung in einem Labor.

Bei *Evernote* – einer Notiz-Applikation u.a. für Smartphones – ist es nun möglich, schnell ein Foto von irgendetwas zu machen, an das man sich erinnern möchte. Irgendein Produkt, das man auch haben möchte; irgendeine visuelle Seltsamkeit, die man sich einsammelt, oder schlicht ein Gesicht: Nun, ist ein Foto erst gemacht, bietet die App zwei Symbole an, mit Hilfe derer man entscheiden kann, wie weiter verfahren wird: ein x – für verwerfen (und einem neuen Versuch) und ein Hackerl - für akzeptieren. Es ist also sogar *nachdem das Bild schon gemacht worden ist* sein Existenzstatus noch in der Schwebe. Und erst, wenn das Hackerl betoucht ist, darf es sein. Bis es, vielleicht schon kurze Zeit später, seine Funktion erfüllt hat.

Dies verweist außerdem auf ein Phänomen, das mensch als *Instabilität* von digitalen Bildern benennen könnte: Nicht nur ihr Existenzstatus schlechthin ist im unmittelbaren Produktionskontext zweifelhaft – auch sein Bildinhalt ist im Fotografiervorgang selbst nicht abgeschlossen: Es gibt Applikationen, die das Bild sofort mit diversen Scharf-, Weichzeichnern, Filtern bearbeiten; sofortigst nachbelichten; es gibt sogar Möglichkeiten, wie das Bild im Nachhinein einen anderen Fokus, eine andere Tiefenschärfe bekommt. Und natürlich kann es mit der Hilfe von Photoshop et.al. so radikal umgestaltet werden, dass eine Ähnlichkeitsbeziehung zur ursprünglichen visuellen Referenz nur mehr mit äußerster Mühe und Wohlwollen hergestellt werden kann- dazu muss mensch sich bekanntermaßen nur diverse Magazincovers zu Gemüte führen.

Wiewohl es auch im 'analogen' Zeitalter durchaus Möglichkeiten der Bildretusche gab, gibt es hier allerdings einen so großen Unterschied in ihrer Qualität und Quantität, sodass die ständige Frage nach der Authentizität des Bildes, seiner 'Echtheit' wohl im Modus des Generalverdachtetes in radikalierter Form aufgeworfen ist – und zwar insbesondere bei öffentlichen Bildern.

Numerische Bilder

Der Produktionsprozess kann sich von der Repräsentation der empirischen Welt prinzipiell loslösen (vgl. Rötzer 1991); d.h. es muss keine notwendige Analogie mehr zwischen

Repräsentierendem und Repräsentiertem existieren, nicht, wie es bei ‚Photoshop‘-Bildern – in verschiedenen Graden – noch gegeben ist: Das Bild kann von Grund auf konstruiert werden, z.B. von einem Computer nach einem bestimmten Algorithmus generiert; vgl. z.B. fraktale Bilder.

U.a. diese Gegebenheit wurde in den 90er- Jahren – wie bei Rötzer - eher noch als Autonomie des Virtuellen oder Immateriellen vom Materiellen aufgefasst, oder gar - wie bei Baudrillard (Baudrillard 1978) - als Verschwinden des ‚Materiellen‘ oder ‚Realen‘ im Simulakrum. Dies ist m.E. durch neuere Entwicklungen – insbesondere im Sektor der sog. *Augmented Reality* – in Frage zu stellen: hier gibt es, wie ich später noch ausführen möchte, eher zunehmende, vielschichtige *Hybridisierungen* zwischen Materialität, Bildlichkeit und Virtualität zu verzeichnen.

Visuelle Technologien

Ich werde nun in den nächsten Schritten versuchen, über die Thematisierung und Kontextualisierung des Visuellen im *Technischen*, über eine Kritik einer relativ allgemeinen Definition von Merzoeff schließlich zu einem Arbeitsverständnis von *Neuen Visuellen Technologien* zu kommen und im Anschluss eine besondere Betonung auf solche zu geben, die die Form und Ausgabe von Bildern ganz wesentlich *Algorithmen* verdanken.

Wie definiert Merzoeff nun Visuelle Technologien?

„By visual technology, I mean any form of apparatus designed either to be looked at or to enhance natural vision...“ (Mirzoeff 2002, 3)

Diese visuelle Technologie ist nun eingebettet in eine Visuelle Kultur. Diese „...is concerned with visual events in which information, meaning or pleasure is sought by the consumer in an interface with visual technology.“ (ebd)

Damit sind nun Ölgemälde genauso gemeint wie Fernsehen oder Internet.

Kritik des Visuellen: Visuelle Technologien als ‚Mixed‘ Technologien

Dies lässt sich nun m.E. über eine Kritik W.J.T. Mitchells an dem Begriff der *visual media* analog kritisieren. Er schreibt: „All media are, from the standpoint of sensory modality, mixed media.“, (Mitchell, 7) Er verdeutlicht das eindringlich am Beispiel der ‚handgemalten‘ Malerei: „...what is the perception of the painting as handmade if not a recognition that a non-visual sense is encoded, manifested and indicated in every detail of its material existence?“ (ebd, 8) und also kann man sagen: „Seeing painting is seeing touching.“ (ebd, 9) Mitchell macht hier die Beobachtung; „...one sense seems to activate or lead to another.“ – in Form der *Synästhesie*, aber auch in Form einer Aktivierung einer ‚Ansicht‘, oder auch des Auditiven (Subvokalen) z.B. durch das geschriebene Wort (vgl. ebd, 11) Mitchell möchte in seinem Text auch eher ‚das Visuelle‘ als zu Hinterfragendes, zu Erklärendes hinstellen, und weniger als etwas, das erklärendes Fundament sein muss für ein ganzes akademisches Feld. (ebd. 13) Das ‚Visuelle‘ an der visuellen Technologie kann man nun m.E. nach in der selben Manier als im Zentrum der Analyse stehendes Explanandum betrachten; insbesondere wenn es um die ‚neuen‘ Varianten geht, die z.B. mit *Touchscreens*² zu tun haben und wo das ‚Visuelle‘ genau etwas ist, das mit tatkräftiger Hilfe von automatisierten Skripten hergestellt wird, die alles andere als transparent zu sein scheinen. Eine etwas undurchsichtige Visualität. Aber dazu später mehr.

*Technik ist stabilisierte Gesellschaft*³

Nun zum zweiten Teil des Wortverbundes der visuellen Technologie: Dazu möchte ich einen aktuellen Ansatz aus der *STS* (Science and Technology Studies) mit dem inzwischen relativ prominenten Namen *ANT* (Actor-Network-Theory) heranziehen und die Definition von Merzoeff ein wenig komplizieren, um sie hernach – modifiziert - umso besser, d.h. adäquater mit algorithmischen Tendenzen im Visuellen (und anderwertig Sinnlichen) auseinandersetzen zu können.

² Früher hats ja von der Mama her als Verhaltensregel des Schauens immer geheißen: „Nicht mit den Fingern oder Händen sieht man, sondern mit den Augen!“ Also museal/bürgerliches Distanznehmen zur *Show*. Heutzutage muss man – Stichwort *Touchscreen* – ja gerade hinfassen, um überhaupt etwas sehen zu können! Eine Paradigmenumkehr: Hand aufs Bild.

³ So der Titel eines Textes von Latour (2006)

Um es kurz zu machen: ANT behauptet: Nicht nur Menschen formen Technologien, die – als Medium - wiederum auf andere Menschen einwirken (= Sozialdeterminismus); auch der umgekehrte Fall ist nicht korrekt: (unausweichliche) Technologien formen die Menschen (=technischer Determinismus); nein: es ist *beides gemeinsam*: technische Artefakte und Menschen sind beide mit *Agency* ausgestattet und agieren als einander assoziierte Handlungsketten (-> Netzwerk). Sog. menschliche und nicht-menschliche Entitäten sind also *analytisch* weitgehend gleichgestellt und *inskripierten* einander Handlungen bzw. Handlungsmodi. Dies möchte ich mit dem Konzept des *Skripts* der französischen Soziologin Madelein Akrich – unter Beihilfe einiger Konzepte und Begriffe von Bruno Latour - verdeutlichen.

Akrich fasst >>Skript<< (oder >>Szenario<<) als materiale Definitionsarbeit eines Designers/ einer Designerin an den Charakteristika seiner/ihrer Objekte, desgleichen aber auch der Akteure, die damit in bestimmter - also hier *inskripiertes* - Weise umgehen sollen.

„Die technische Realisierung der Vorstellungen des Innovators [e.g. Designers, SC] über die Beziehungen zwischen einem Objekt und den es umgebenden Akteuren ist also ein Versuch der Vorherbestimmung der Settings, die Benutzer sich für ein bestimmtes Stück Technik vorstellen sollen, und die Präskriptionen (Notizen, Verträge, Ratschläge usw.), die es begleiten.“

(...) definieren technische Objekte wie ein Filmskript den Rahmen einer Handlung zusammen mit den Akteuren und dem Raum, in dem sie agieren sollen.“ (Akrich 2006, 411)

Zu diesem Zweck werden in das technische Objekt *Handlungsprogramme* eingebaut (inskripiert), die mögliche *Antiprogramme* (z.B. Ignoranz desselben; Zweckentfremdungen etc.) von Usern desselben antizipieren und mit in Rechnung stellen. (vgl. Latour 2006, 373) Welche Handlungen sich nun aus dem – bei Erfolg stabilisierten – Ding/Mensch-Arrangement ergeben, ist schließlich Sache des/der TechniksoziologIn - zumindest um sie nachzuzeichnen. Es wird davon ausgegangen, dass die Intentionen (Skripte), die am Anfang in das Objekt einfließen, nicht mehr dieselben sind, nachdem sie sich mit den verschiedenen, handelnden Akteuren verwoben haben. Ein Mensch, der Dinge verwendet, wird in seinen Handlungen letztendlich auch von den eingeschriebenen Handlungsprogrammen dieser Dinge geformt (die ja auch in der Folge der Antiprogramme durch andere - wenn auch womöglich zeitlich versetzt - Mithandelnde modifiziert sind.) und ebenso die Skripte der

jeweiligen Technologien.

Nehmen wir als Beispiel eine analoge Kleinbildkamera als visuelle Technologie im Sinne von ANT (und Mitchell), wobei ich hier mehr im Sinne einer eher skizzenhaften Demonstration der Denklogik nur einige wenige Elemente nennen möchte und insbesondere auch von der äußerst komplexen Technik- und Produktionsgeschichte absehen werde.

Die analoge Standard - Kleinbildkamera – und ich sehe auch einmal davon ab, ob Kompakt oder Spiegelreflexkamera – ist erst einmal so designt, dass der/ UserIn dazu angehalten ist, sie in beide Hände zu nehmen (es gibt dafür auch Griffe bzw. entsprechende Buchtungen am Gehäuse), wobei die möglichen (natürlich vorprogrammierten) Einstellungen und der Auslöser sich in Reichweite der Finger der rechten Hand befinden: es ist hier also der/die *imaginierte UserIn* einmal – sieht man von Sonderanfertigungen ab - als RechtshänderIn gedacht.

Die möglichen Einstellungen erlauben nicht kontinuierliche, sondern bestimmte, disparat wählbare Blendenöffnungsgrößen und –zeiten, die auch automatisiert gewählt sein können. Desgleichen die Schärfeneinstellung am Objektiv vorne: die verbreitetsten Modelle hatten spätestens in den 90ern Autofokus; d.h. der Apparat setzte den Schärfepunkt des Bildes nach den ihm eingebauten Kriterien. Auch die mögliche Helle des Bildes – bzw. die damit zusammenhängende kleinst-mögliche Blendenöffnungszeit ist durch die Lichtempfindlichkeit des Objektivs – und des Filmes vorgeprägt.

Der Apparat ‚übersetzt‘ außerdem eine bestimmte Verbildlichungsabsicht eines/r wohl menschlichen Fotografierers/in in ein Bild unter Beteiligung einiger hier zu beleuchtender Momente, die das *Skript* des Verbildlichungsprozesses nunmehr soweit vervollständigen wie es hier möglich ist:

Es stellt das mit eigenen Augen Gesehene in einen Rahmen hinein. Welcher Rahmeninhalt das ist, ist dem/r UserIn bzw. seinen/ ihren kulturellen Kompetenzen überlassen, e.g.

Einschreibungen in die Akteure als praktisches Wissen, wie dies technische Ding denn sich geriert, wenn es so oder eben so befangert wird. Aber *dass* ein Rahmen - im in der Kamera eingebauten Format – gewählt werden muss, ist vordesignt. Ebenso die von der Optik kommende Schärfe/Unschärfe: So etwas gibt es im ‚natürlichen‘ Blick nicht. Und natürlich die Zweidimensionalität, eine bestimmte Farbgebung; die über die mögliche Lichtempfindlichkeit bestimmte möglichen Orte und Zeiten des Fotografierens (das ändert sich bis zu einem

gewissen Grad mit dem Blitz); aber auch die Möglichkeit, zuvor Unsichtbares zu verbildlichen: festgefrorene Bewegungen im Mikrosekundenbereich; durch Teleobjektive optisch vergrößerte Dinge oder Sterne, Planeten; in Zusammenschaltung mit Mikroskopen unfassbar kleine Dinge usw.

Kurz und etwas flapsig: Durch den (analogen) Fotoapparat wurde ein anderer Blick möglich; und dieser Blick, diese andere Art der Wahrnehmung, würde wohl die Actor-Network-Theory argumentieren, schreibt sich auch in die Wahrnehmungsschemata der menschlichen Akteure ein. Und zwar – auch über die *Praxis* des Fotografierens – unter Berücksichtigung der Inskriptionen des Fotoapparats.

Zusammenfassend

Visuelle Technologie muss m.E.

1. immer als *Mixed Technology* gedacht sein; d.h. die *performative* Referenzierung auf andere Sinne denn jenen des Sehens soll – ganz im Sinne von Mitchell – Reflexionen darüber anstoßen, wie sich denn nun Visualität im Erregen und Ausgreifen unserer Sinne hineingliedert.

2. besteht, wie ich mit Hilfe von ANT zeigen wollte, eine Wechselbeziehung von ‚menschlichen‘ und ‚technischen‘ Akteuren, die sich nicht allein im Formen des einen durch den anderen erschöpft, sondern *symmetrisch* erfolgt. Damit möchte ich auch darauf hinweisen, dass es sich wohl lohnt, in Zeiten zunehmender Technisierung des Visuellen die Prozesse der Technisierung en Detail zu begutachten, um m.E. überhaupt Aussagen über die darin entstehende Bildlichkeit machen zu können.

2. Teil: Neue Visuelle Technologien

In einer Spezialausgabe der Zeitschrift *Visual Studies* des Jahres 2011⁴ wird mit der Überschrift: „*New visual technologies: shifting boundaries, shared moments*“ über aktuelle Verschiebungen und Neuorientierungen im Feld des Visuellen nachgedacht. Im Titel stecken schon Begriff und ein Großteils der hier angewandten Arbeitsdefinition.

⁴Visual Studies, Vol.26, No.2, June 2011

Der Begriff wird recht allgemein gehandelt; als „...socio-technical *things* embedded and emplaced in the world.“ (Graham 2011, 87) – und in dieser Allgemeinheit durchaus im selben epistemologischen Kontinuum mit der Actor-Network-Theory –wiewohl die Art der Verbindungen von Mensch und Ding hier nicht angesprochen sind; Graham et.al unterscheiden vier Typen von Neuen Visuellen Technologien: carriers – wie das Internet; production technologies – wie Kameras und Mobiltelefone; display technologies - wie Öffentliche Bildschirme oder mobile Projektoren und Services – wie Weblogs oder Flickr. (Vgl. ebd.)

Die *Shifting Boundaries* – sich verschiebende Grenzen – beziehen sich auf unzählige Grenzen und Abgrenzungen des alltäglichen Lebens, die sich mit und durch die neuen Technologien verschöben: Der Unterschied zwischen Arbeit und Freizeit (oder *work* und *home*), der damit verwandte Unterschied zwischen Intimsphäre und öffentlicher Sphäre; zwischen dem physikalischen und dem digitalen; Amateurementum und Professionalismus usw. usf. (Vgl. ebd, 88); *Shared Moments* bezieht sich auf synchrone soziale Vernetzungen, die über das Netz möglich sind; ein Phänomen, das ich im ersten Teil schon ein wenig diskutiert habe.

Dies alles sind ohne Frage wichtige und zentrale kulturwissenschaftliche bzw. kultursoziologische Problemstellungen; ich möchte aber den Fokus – noch etwas skizzenhaft und etwas soziologischer – auf mögliche Folgen für eine Handlungstheorie legen: wie sich *Agency* in all diesen oben genannten Programmen und Dimensionen hin zu automatischen Algorithmen verschiebt; und dabei besondere Obacht auf einen bestimmten –m.E. nach entscheidenden Programmkomplex legen.

Zu diesem Behufe werde ich mich mit einem Text von William Uricchio beschäftigen. Dieser nun verlegt den Schwerpunkt einer Analyse neuer visueller Technologien ins Fragen nach den „...algorithmically defined relations between the viewing subject and the world viewed...“ (Uricchio 2011, 25).

Algorithmic Turn

Uricchio argumentiert in diesem Text die seiner Meinung nach herausragende Rolle von *algorithmischen Interventionen* zwischen dem wahrnehmenden Subjekt und dem

dargestellten Objekt in einer Weise, die ihn zur Annahme drängt, dies sei ein radikaler Wandel vom „Welt-bild“ des Martin Heidegger hin zu einem neuen visuellen Regime. Einem Regime, in dem Algorithmen – das sind allgemein verstandene endliche Sequenzen von Anweisungen, die wiederum ein Computerprogramm automatisiert ausführt – entscheidende Handlungsrelevanz haben.

Uricchios These: „My argument in a nutshell is that over the past decade or so we have had increased access to new ways of representing and seeing the world, ways dependent on algorithmic interventions between the viewing subject and the object viewed.“ (Uricchio 2011, 25)

Uricchio verdeutlicht diese seine These nun an zwei Beispielen neuer visueller Technologien: Das in Kooperation von Microsoft und der University of Washington entstandende Bildprogramm *Photosynth*, das ich relativ kursorisch betrachte und die Klasse der *Augmented Reality* – Applikationen, der ich mich danach im Besonderen zuwenden werde.

Photosynth

Photosynth⁵ ist ein Programm, das mittels eines Algorithmus befähigt ist, verschiedene Bilder auf ihren gemeinsamen *geographischen* Bezug hin abzutasten – nachdem sie vom User/ der Userin als von einem gemeinsam Ort stammend markiert oder *getaggt* sind. Dazu generiert die Applikation eine sog. Punktwolke (Point Cloud), die diese geographische Umgebung dreidimensional anhand bestimmter Kriterien (Kanten, Ecken, heraustechende Kontraste etc) markiert; diese Punktwolke wird wiederum sonach als quasi-objektiver Bezugspunkt für die Art und Weise der Zusammensetzung dieser unterschiedlichen Bilder in unterschiedlichen Perspektiven von einem Ort verwendet. Die so – bis zu 600 – zusammengesetzten Bilder eines Ortes können vom/von der UserIn nach der algorithmischen Intervention als relativ nahtlos ineinander-übergend betrachtet werden, wobei sich folgende Differenzen zum klassischen, seit der Renaissance geübten Perspektivenbild ergeben:

- Die Bilder sind von unterschiedlichen Menschen gemacht; zu unterschiedlichen Zeiten; in unterschiedlichen Perspektiven; mit unterschiedlichen Kameras, in unterschiedlichen

⁵www.photosynth.net; Beispielhaft: <http://photosynth.net/view.aspx?cid=a2eb820f-79d1-429c-86ee-5537ddd48d75>, Zugriff: 5.5.2013, 10.00

Auflösungen, Tiefenschärfen etc.pp

- Ein für die UserIn *intransparentes* Software-Skript selektiert die Bilder: z.B. werden keine Nachtbilder akzeptiert; Bilder von Bäumen sind schwierig, weil die sich durch Wind und Wetter ständig ändern und nicht mehr als gleiche erkannt werden; landschaftliche Differenzen zur etablierten Punktwolke, die über die Zeit auftreten können, werden automatisch ausgeschlossen; etc. (vgl. ebd., 29)

- Der/die BetrachterIn sieht also ein multi-perspektivisches Bild von multiplen AutorInnen visualisiert, und von einem Algorithmus in bestimmter Weise zusammengesetzt.

Es wird also das klassische Subjekt-Objekt Verhältnis (des Bildes) durchkreuzt.

„Authorship, in this context, is both problematic and pluriform-. Although mostly effaced, it is the author of the individual photographs; largely enacted, it is the author of the experience – that is, the navigating user; fundamentally enabling, it is the author of the algorithm; and in terms of what we actually see and select from, it is the algorithm as author.“ (Uricchio, 31)

Augmented Reality

Augmented Reality – ‚erweiterte‘ Realität – bezeichnet eine Klasse von Software-Applikationen, die auf ein durch ein Kamera/Bildschirm- System bereitgestelltes Bild aktuellen Geschehens eine zweite, virtuelle Ebene überblendet oder überlagert (overlay), die die visualisierte empirische Welt in bestimmter Weise markiert und/oder beschreibt.

Uricchio beschreibt dabei drei unterschiedliche Typen, das ‚Reale‘ und das ‚Virtuelle‘ zu verbinden. (Vgl. ebd, 31)

- *fiduciary markers* (= ca. Markierung durch Referenzpunkte): sind die einfachsten:

Kodifizierte graphische Tags (Schilder) – vergleichbar mit einem Strichcode – initialisieren, wenn durch die Applikation erfasst, ein graphisches Overlay, das allerdings nur genau an diesem (geographischen) Ort, mit genau jenem Schild im Bild, funktioniert.

- *Digital Compass Tracking Systems* (z.B. wikitude, layar) verwenden das in Smartphones und vielen Kameras eingebaute GPS-System und die Analyse der Distanz zu WLAN-Netzwerken für die Lokalisierung; einen eingebauten Kompass für die Erkennung der

Blickrichtung und einen Beschleunigungsmesser, um das im jeweiligen Bild gerade erkennbare so indirekt zu identifizieren, und daraufhin ein Orts- und Perspektiven-abgestimmtes Overlay zu legen.

- *Natural Feature Tracking Systems*: die aufwendigste Form der AR; die Applikation hat eine eingebaute automatische visuelle Erkennung und kann verschiedene Datenebenen damit verbinden und überblenden – und steht damit z.B. in technischer Verbindung mit der automatischen Gesichtserkennung von Videoüberwachung; und mit Ansätzen zur Konstruktion halb- und schließlich vollautonomer Waffen – in nicht allzuferner Zukunft⁶.

Diese visuellen Erkennungssysteme – gleich welchen obigen Typs – sind aber immer noch größtenteils davon abhängig, dass man die Kamera in bestimmtem Winkel zum zu erkennenden Objekt/Subjekt hält, den das System für die weitere Verarbeitung akzeptiert; das AR-System gibt hier also auch bestimmte Körper- bzw. Handhaltungen vor, um verwendet werden zu können.

Weiter bieten AR-Systeme eine *embodied* Perspektive – sie ist immer weitgehend deckungsgleich mit der Perspektive des/der UserInn: also auch hier ein Abgehen von der alten visuellen Konvention der ‚autoritativen‘ Perspektive, die vom Betrachter/der BetrachterIn als Gegeben hingenommen werden muss.

Photosynth und AR-Systeme bergen Unterschiede: auf der einen Seite unzählige Perspektiven verschiedener Zeiten, auf der anderen die verkörperte Perspektive in ‚Echtzeit‘: auf der einen Seite Bewegung *im* Bild (Photosynth: beim Übergang von einem der Bilder zum anderen, der meist auch einen Perspektivenwechsel beinhaltet), auf der anderen Bewegung *mit* dem Bild (Vgl. ebd, 32); doch beide Programme bzw. Programmklassen „...share a fundamental realignment of subject-object relations thanks to their algorithmic processing layers. (...) the algorithmic domain ultimately determines *what* we see, and even *how* we see it.“ (ebd) Das – so Uricchio – heißt letztendlich: „AR systems effectively overlay the viewer’s access to the physical world with specific (and selectable) grids of signification.“ (ebd., 33)

⁶ Vgl. dazu einen Report von Human Rights Watch: „Losing Humanity – The Case Against Killer Robots“: <http://www.hrw.org/reports/2012/11/19/losing-humanity>, 30.04.2013, 10.00

Zwei Beispiele

Mitsubishi Electric MeView Augmented Reality Maintenance Enterprise Application

Hier handelt es sich um eine AR – Bedienungsanleitung von Mitsubishi zur Wartung einer Klimaanlage⁷. Dies ist ein Beispiel für den dritten und am höchsten entwickelten Typ in Uricchios AR-Typologie; ich möchte über eine kleine Skizze dieser Applikation eine Ahnung von dem vermitteln, was aller Voraussicht nach in naher Zukunft eine von vielen ähnlichen in den Alltag integrierten ‘Hilfs’programmen darstellt. Dabei beschreibe ich die App hauptsächlich anhand des oben verlinkten Videos.

Die Handlungsanweisungen zur Wartung werden – im Unterschied zu einer Bedienungsanleitung aus Papier - *am Ding selbst* gezeigt; und das Ding in in ihm inskripierten – und durch die AR-Applikation graphisch prä-inskripierten - vom menschlichen Akteur zu erreichenden Endzustand.

Vielleicht soll hier auch weniger von einer Prä-Inschrift die Rede sein – die ja bei Akrich so etwas wie Vorschriften oder Begleitschriften darstellen, sondern einem *in actu* ausbuchstabilerten, auf direktestem Wege demonstrierten und demonstrierenden digitalen Skripts (ein ‘hybrid-simulatives Live-Skript’?), das so die ‘richtige’ Handhabung respektive in diesem Fall: Wartung des Dings weiter stabilisiert und die Wahrscheinlichkeit weiter erhöht, dass der menschliche Akteur die vom Design vorgesehenen Handgriffe auch genau so vollzieht, wie es von ihm/ihr auch erwartet wird. Dies birgt einerseits die Möglichkeit der Reduzierung menschlicher kreativer Handlung (auch betreffend einer möglichen Zweckentfremdung einer Technologie; oder einfach andere Wege der Bedienung, die ja immer möglich sind) – andererseits vielleicht sogar auch so etwas wie ‘Empowerment’: immerhin kann nun auch der Laie ihm oder ihr fremde Dinge bedienen oder reparieren; und braucht nicht extra eine/n Fachmann/frau zu Hilfe ziehen.

Vielleicht wäre diese Stabilisierung der inskripierten Ding-Be-handlungserwartung auch auf andere Bereiche von Neuen Visuellen Technologien anwendbar: es werden z.B. durch ins Gesichtsfeld eingeblendete Navigationsanweisungen –in verstärktem Maße noch als in bisher vorhandenen Formen im Auto oder Smartphone – die vorgeschriebenen bzw eingeschriebenen Wege eher befahren, als jene, die in der jeweiligen Datenbank nicht

⁷ http://www.youtube.com/watch?v=iz4ykMn3UR4&list=PLaQauau23m3OyK7Wdk70a_mgon_CwH3RZ, 30.4.2013, 15.30

verzeichnet sind⁸ oder als solche verzeichnet sind, die ökonomischen Effizienzkriterien nicht genügen und so fort, je nach Programmierung.

Des Weiteren werden durch AR – Technologie in einem solchen visuellen Umgebungsfluss bestimmte Dinge oder Prozesse hervorgehoben, *markiert*, und andere nicht. Es ist hier vielleicht ein Analogon einer Kritik an Standardkarten anwendbar: Karten, in denen ja immer auch zuvörderst staatliche und ökonomisch wichtige – und also hegemoniale und insbesondere *quantifizierte*⁹ – Elemente –herausgehoben sind und andere nicht. Dieses Orientieren und sich Bewegen nach – natürlich im Detail nachzuweisenden - hegemonialen Linien wäre im Übrigen auch das genau Gegenteil des *derive* (herumschweifen) der französischen Situationisten – einem im assoziativen Streifen, Herumschweifen in der Stadt, der aktiven Konstruktion einer individuellen 'Psychogeographie' (vgl. Ohrt 1995).

Allerdings wird eine Übertragung bzw. Übersetzung der realräumlichen Machtverhältnisse ins Virtuelle bzw. ins Hybride-Augmentierte möglicherweise durch die darin auch geographisch sichtbar gemachten Kontakte der eigenen Social Media-Mitgliedschaften relativiert. Auch das wird sich in der Praxis zeigen.

Google Glass

Google Glass ist ein Projekt von Google¹⁰, das sich mittlerweile im Endstadium der Entwicklung befindet und nunmehr seit einigen Wochen auch von Privaten getestet wird. Dazu gibt es und gab es auch im Vorfeld massive Kontroversen zu diesem Stück Technologie. Und damit ist es auch ein großartiges Beispiel *in status nascendi* für die Konstruktion und Etablierung einer innovativen Technologie und ein gutes Beispiel für eine Grenzverschiebung von Graham et.al: Gerade der für Außenstehende nicht eruierebare Aufnahmestatus der integrierten Kamera suggeriert ständige visuelle Überwachung – und möglicherweise einen Live-Upstream ins Netz; sodass private Situationen ständig Gefahr laufen, gerade als öffentliches Bild gesendet zu werden. In Reaktionen darauf ist Google Glass schon einmal in einigen Lokalen verboten worden.¹¹

⁸ Google Maps statt Mental Maps!

⁹ Das Phänomen der Quantifizierung sämtlicher Umgebungen und sogar Lebensumstände müsste noch separat behandelt werden. Insbesondere auch der Status der hier betrachteten Technologie in der sogenannten ‚Quantify-yourself‘ ‚Bewegungen‘, vgl. z.B. <http://quantifiedself.com>, 5.5.2013, 12:00; eine Homepage mit dem Untertitel: „self knowledge through numbers“

¹⁰ <http://www.google.com/glass/start>, Zugriff: 05.05.2013, 11:00

¹¹ Vgl: <http://www.zeit.de/digital/mobil/2013-05/google-glass-verbotten>, 15.05.2013, 10:00

Google Glass ist eine Brille mit am rechten Auge eingeblendeten Bildschirm, einer Kamera, Mikrofon und einem haptischen Bedienelement in Form des rechten Ohrenbügels – mit einem integrierten Minicomputer. Sie ist mit ähnlicher Funktionalität ausgestattet wie ein Android-Smartphone, der entscheidende Unterschied liegt natürlich in der permanenten Verfügbarkeit von visuellen Informationen; von Über- und Einblendungen ins 'natürliche' Gesichtsfeld, also sozusagen eine 'full-embodied' Perspektive.

Eine solche Brille zusammen mit der Möglichkeit von automatischer visueller Erkennung wäre die maximale Variante von Augmented Reality, und auf diesem Gebiet wird natürlich kräftig geforscht: Eine neue App mit dem Namen 'Insight' soll z.B. die Fähigkeit haben, Menschen nicht am Gesicht, sondern an ihrem Kleidungsstil mit Hilfe eines 'visual fingerprint' zu erkennen.¹²

Ausblick und Schluss

Google Glass stellt also die 'höchste' Stufe des *mediatisierten Sehens* dar: Ein Sehen, das nicht nur mittels aufbereiteter, standardisierter und normalisierender digitaler Visualitäten mitdirigiert wird, sondern nicht zuletzt durch automatische Erkennungsprogramme, die das Sehen/Erkennen teilweise für den Menschen erledigen; wo also *die kognitiven Prozesse des Alltagswahrnehmens* zum Teil in Maschinen ausgelagert sind; Maschinen, die nach Regeln in dieses Sehen intervenieren, die den Menschen wenig bis gar nicht bekannt sind; und die ihr – in der Folge hybrides - Handeln in der Folge also auch aufgrund von maschinellen Wahrnehmungen formen. Wenn hier nicht Aufklärungsarbeit geschieht und die hier (mit)entscheidenden Algorithmen für den/die UserIn transparent gemacht werden, kommt es wohl oder übel zu einem großen Schritt weg vom aufklärerischen und demokratischen Ideal der Mündigkeit und hin zu einer – durch immer gleiche Algorithmen ermöglichten– Normalisierung des Alltagshandelns.

Es handelt sich hier um eine Technologie, die über ihre Vorstrukturierungen hochgradig handlungsrelevant ist; diese Relevanz ist aber nicht transparent, nicht ersichtlich.

Diese Asymmetrie ist eine wichtige politische Implikation.

¹² vgl. <http://www.newscientist.com/article/mg21729075.600-google-glass-app-identifies-you-by-your-fashion-sense.html> und <http://synrg.ee.duke.edu/papers/insight-final.pdf>, Zugriff jeweils am 30.04.2013, 15.00

Ebenso hat dies meiner Ansicht nach erhebliche Relevanz auch für eine soziologische Bildtheorie:

“Nahezu alle bildtheoretischen Ansätze stimmen darin überein, dass bildlicher Sinn aus einem Zusammenspiel zwischen der Wahrnehmung der Materialität des Bildträgers, der appräsentativen und imaginativen Wahrnehmung des im Bild Dargestellten einschließlich der Art und Weise seiner Darstellung, dem Stil oder der Form, sowie dem wahrnehmenden Subjekt/Beobachter/Bildproduzent in einer mindestens triadischen Bedeutungsrelation entsteht.”, (Breckner 2010, 114)

Vielleicht muss mensch – folgt eben dieser den weiter oben gemachten Ausführungen - möglicherweise so etwas wie automatisierte Skripte als mitformend für die „.. Ebene des im Bild Dargestellten einschließlich der Art und Weise seiner Darstellung..“ in Rechnung stellen. Und mit der Behauptung der Relevanz dieser Beobachtung müsste mensch folglich verstärkt so etwas wie eine Techniksoziologie mit ihren wohletablierten analytischen Instrumentarien in die Erforschung des Visuellen mitintegrieren – um die Herausforderungen durch eine algorithmisch mitgeformte menschliche Gesellschaft bewältigen zu können.¹³

Ich hoffe, ich habe das mit dieser kleinen Arbeit– zumindest in skizzenhafter Form – ein wenig andeuten können.

Literatur

Akrich, Madeleine: *Die De-Skription technischer Objekte*, in: Belliger, Andréa; Krieger, David J. (Hg) - *ANThology – Ein einführendes Handbuch zur Akteur-Netzwerk Theorie*, transcript, Bielefeld 2006, 407-428

Baudrillard, Jean: *Agonie des Realen*, Merve, Berlin 1978

¹³ Daniel Suarez – Autor von Tech- Thrillern wie *Daemon* oder *Kill Decision*, die für ihre technische Genauigkeit vielfach gelobt werden, formuliert dies in einem FAZ – Interview recht drastisch: „...Algorithmen [sind]die Gesetze des 21. Jahrhunderts - die zahllosen gehorsamen Vollstrecker des wuchernden hierarchischen sozialen Systems. Sie gleichen unserem offiziellen gesetzlichen Rahmen und erlangen zunehmend eine ähnliche Macht über uns. Algorithmen kodifizieren den Willen ihrer Schöpfer. Indem sie Regeln und Parameter festlegen, begrenzen sie den menschlichen Einfluss, strukturieren unsere Interaktionen, kanalisieren das Spektrum der Möglichkeiten in eine endliche Zahl von Pfaden.“, <http://www.faz.net/aktuell/feuilleton/debatten/digitales-denken/gesprach-mit-daniel-suarez-wir-werden-mit-system-erobert-14089.html>, Zugriff: 30.3.2013, 10.00

Breckner, Roswitha: *Sozialtheorie des Bildes – Zur interpretativen Analyse von Bildern und Fotografien*, Transcript, Bielefeld 2010

Graham , Connor et. al (Hg): *New visual technologies: shifting boundaries, shared moments*, Visual Studies, 2011, 26:2, 87-91

Lapenta, Francesco: *Locative media and the digital visualisation of space, place and information*, Visual Studies, 2011, 26:1, 1-3

Latour, Bruno: *Technik ist stabilisierte Gesellschaft*, in: Belliger, Andréa; Krieger, David J. (Hg) - *ANTHology – Ein einführendes Handbuch zur Akteur-Netzwerk Theorie*, transcript, Bielefeld 2006, 369-399

Mirzoeff, Nicholas: *What is visual culture?*, in Mirzoeff, Nicholas (Hg): *The Visual Culture Reader* (2nd ed), London 2002, 3-13

Mitchell, W.J.T.: *There are no visual media*, in Mirzoeff, Nicholas (Hg): *The Visual Culture Reader, 3rd edition*, Routledge, New York 2013, 7-14

Ohr, Roberto (Hrsg.): *Der Beginn einer Epoche. Texte der Situationisten*. Edition Nautilus, Hamburg 1995,

Rötzer, Florian: *Mediales und Digitales – Zerstreute Bemerkungen und Hinweise eines irritierten informationsverarbeitenden Systems*, in: Rötzer, Florian (Hg): *Digitaler Schein – Ästhetik der elektronischen Medien*, Suhrkamp, Frankfurt/Main 1991, 9-81

Uricchio, William: *The algorithmic turn: photosynth, augmented reality and the changing implications of the image*, Visual Studies, 2011, 26:1, 25-35

Virilio, Paul/ Kittler, Friedrich: *Die Informationsbombe: Paul Virilio und Friedrich Kittler im Gespräch*: <http://paedpsych.jk.uni-linz.ac.at:4711/LEHRTEXTE/KittlerVirilio.html>, Zugriff am 30.4.2013, 12:00