Algorithmisches Panopticon

Algorithmen erfassen, analysieren und beurteilen jede Regung im öffentlichen Raum einer Großstadt.
Was wie Science-Fiction klingt, wird von Wissenschaft und Forschung längst mit Hochdruck für einen baldigen Einsatz vorangetrieben. Angestrebt wird eine **möglichst umfassend automatisierte Überwachung**, die nicht nur die Aufgabe der OperateurInnen übernimmt, sondern alle Möglichkeiten digitaler Datenerhebung und Informationsverarbeitung ausreizt. Dabei wird jedoch konsequent ignoriert, wie wichtig es für einen mit den Grundrechten vereinbaren Einsatz ist, die erhofften Vorteile gegen mögliche Risiken abzuwägen.

Mit seiner als Buch überarbeiteten Diplomarbeit aus dem Gebiet Informatik und Gesellschaft, die auch für NichttechnikerInnen verständlich ist, geht Benjamin Kees einen ersten Schritt der versäumten Technikfolgeabschätzung nachzukommen. Zur Identifikation gesellschaftlicher Probleme skizziert er anhand aktueller Forschung und vorhandener Technik ein **zu erwartendes Überwachungskomplettsystem** und untersucht es aus technischer, psychologischer, soziologischer und rechtlicher Perspektive.

Er stößt dabei auf Datenschutzalbträume, rassistische und diskriminierende Algorithmen und entlarvt die Rolle menschlicher OperateurInnen als Feigenblatt einer unwägbaren Vollautomatisierung.

Die Diplomarbeit wurde 2014 vom *Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung* mit dem **FIfF-Studienpreis** für herausragende Qualifikationsarbeiten aus dem Bereich Informatik und Gesellschaft ausgezeichnet:

"Die Arbeit behandelt ein immer noch hochaktuelles, gesellschaftlich und politisch relevantes Thema an der Schnittstelle von Informatik und Gesellschaft. Sie kommt zu wichtigen Ergebnissen für die Debatte um unsere Rechte und unsere Sicherheit."
FIfF e.V.

www.algoropticon.de

Benjamin J. Kees | Algorithmisches Panopticon

Benjamin J. Kees

ALGORITHMISCHES PANOPTICON

Identifikation gesellschaftlicher Probleme automatisierter Videoüberwachung

Auszug aus der Laudatio für den FIfF-Studienpreis 2014

Wenn wir heute durch eine deutsche Großstadt spazieren, schauen hunderte von Kameras auf uns herab. Umfassende Videoüberwachung gilt allenthalben als die Lösung vieler Probleme nicht nur der Kriminalitätsbekämpfung. Und kann die heutige Überwachung eine Tat nicht verhindern, werden flugs Rufe laut, sie weiter zu verstärken und intensivieren. Die Befürworter, die nicht müde werden, die Forderung nach Videoüberwachung immer wieder zu wiederholen, versprechen sich davon einfache Lösungen für Sicherheitsprobleme, leichte Überwachung auch unübersichtlicher Räume, zentraler Betrieb mit wenig Personalaufwand, erschwingliche Technik, ohne dass ihr Betrieb besondere Kenntnisse erfordert. Die praktische Anwendung verweist diese Hoffnungen häufig ins Reich der Mythen. Und: Videoüberwachung ist auch und vor allem ein großes Geschäft, für Entwickler, Hersteller und Betreiber – das wird aber gerne verschwiegen.

Ist der öffentliche Bereich dabei noch einigermaßen kontrolliert, erleben wir im privaten Bereich erheblichen Wildwuchs. Das Ergebnis sind häufig Videoüberwachungsanlagen, die völlig unkontrolliert alles überwachen, was in ihren Fokus kommt, und nicht selten einer rechtlichen Überprüfung nicht standhalten. Dabei ist nicht nur der Nutzen der Überwachungsanlagen fraglich, sie bergen auch erhebliche Risiken für die Menschenrechte. Sie greifen in die Privatsphäre all jener ein, die ihren Sichtbereich durchqueren. Die ständige Möglichkeit der Überwachung beeinflusst unser Verhalten – wie bereits Jeremy Bentham im 18. Jahrhundert mit seinem Konzept des Panoptikon gezeigt hat. Und durch automatisierte Auswertung und Speicherung der Daten auf Vorrat werden umfassende Bewegungsprofile der Menschen möglich.

Nachdem wir heute wissen, dass Geheimdienste alle Daten speichern und auswerten, die sie bekommen können, ist zudem klar: Auch die Daten aus der Videoüberwachung fließen in den großen Pool ein, mit dem die Sicherheitsbehörden ihren Cyberwar gegen uns alle führen.

Diesen Fragen widmet sich die vorliegende Arbeit von Benjamin Kees, die an der Humboldt-Universität zu Berlin entstanden ist. Im Rahmen der Arbeit wird ein automatisiertes System zur Videoüberwachung entworfen, das angesichts der heute bestehenden technologischen Möglichkeiten vorstellbar oder sogar wahrscheinlich ist. Durch das beschriebene, repräsentative Informatiksystem kann die Videoüberwachung und der Kontext ihres Einsatzes anhand eines konkreten Systementwurfs umfassend untersucht werden. Diese Analyse anhand der Anforderungen und des technischen Aufbaus eines Informatiksystems, die über eine rein theoretische Betrachtung hinausgeht, zeichnet die Arbeit aus. Technikfolgen werden anhand der tatsächlich bestehenden Möglichkeiten der Technikanwendung analysiert.

Die Arbeit geht dabei über die unmittelbaren Auswirkungen hinaus, und behandelt auch mittelbare, langfristige Auswirkungen des Systemeinsatzes. Neben den direkten Effekten der Videoüberwachung werden so auch indirekte Effekte und die langfristigen Auswirkungen in die Untersuchung eingebunden. Die Arbeit kommt damit zu sehr umfassenden Ergebnissen.

„Zu Ihrer Sicherheit wird dieser Bereich videoüberwacht", so steht es häufig auf den gesetzlich vorgeschriebenen Warnschildern. Richtig müsste es heißen, das bestätigt die Arbeit eindrucksvoll: „Vorsicht! In diesem Bereich werden Ihre Rechte der Videoüberwachung geopfert!"

Das FIfF hat sich, im Gleichklang mit anderen Initiativen und digitalen Menschenrechtsorganisationen, in der Vergangenheit immer wieder mit den Risiken der Videoüberwachung auseinandergesetzt. In diesem Umfeld ist die Arbeit von Benjamin Kees besonders wertvoll und von hohem Nutzen. Mit ihrer Untersuchung der gesellschaftlichen Auswirkungen der Videoüberwachung behandelt die Arbeit ein immer noch hochaktuelles, gesellschaftlich und politisch relevantes Thema an der Schnittstelle von Informatik und Gesellschaft, das durch die Enthüllungen über die Aktivitäten der Geheimdienste noch einmal an Brisanz gewonnen hat. Sie ist interdiziplinär aufgebaut, umfassend, und kommt zu wichtigen Ergebnissen für die Debatte um unsere Rechte und unsere Sicherheit. Aus diesem Grund hat sich die Jury des FIfF-Studienpreises einhellig für die Auszeichnung der Arbeit entschieden.

Stefan Hügel, Vorstand des FIfF

Benjamin J. Kees

Algorithmisches Panopticon

Identifikation gesellschaftlicher Probleme automatisierter Videoüberwachung

Überarbeitung der Diplomarbeit
Identifikation gesellschaftlicher Probleme automatisierter Videoüberwachung
entstanden bei Prof. Wolfgang Coy am Lehrstuhl für
Informatik in Bildung und Gesellschaft und
Prof. Hartmut Wandke am Lehrstuhl für Ingenieurpsychologie der
Humboldt-Universität zu Berlin

Ich danke Jörg Pohle, Constanze Kurz, Rainer Rehak und André Riefstahl für wegweisende Einflüsse und Beratung. Besonders danke ich Wolfgang Coy und Hartmut Wandke, die an der Humboldt-Universität den für diese Arbeit nötigen Raum geschaffen haben.

Benjamin J. Kees
»*Algorithmisches Panopticon*«
© 2015 der vorliegenden Ausgabe: Edition MV-Wissenschaft
Die Edition MV-Wissenschaft erscheint im
Verlagshaus Monsenstein und Vannerdat OHG Münster
mv-wissenschaft.com

Satz: Benjamin J. Kees und LaTeX
Umschlag: Benjamin J. Kees
Foto Justitia: © Stefan Welz - Fotolia.com
Druck und Bindung: Monsenstein und Vannerdat
ISBN x-xxxxxx-xx-x

Früher hat man dem Computer ein Problem übergeben,
wenn man es verstanden hatte,
heute ist es andersrum.

JOSEPH WEIZENBAUM

Abstract

Videoüberwachung, die trotz nicht nachgewiesener Effektivität und negativer Auswirkungen auf Individuen und Gesellschaft weltweit massiv ausgebaut wurde, soll durch Algorithmisierung in Zukunft so automatisiert wie möglich gestaltet werden. Dazu sollen nicht nur die bisherigen Aufgaben der OperateurInnen weitestgehend übernommen werden, sondern wird auch eine Erhebung, Verarbeitung und Nutzung von Daten angestrebt, die über menschliche Kapazitäten weit hinausgeht. Auf Grund der technischen Funktionsweise, dem grundsätzlichen Charakter von Überwachung und der Automatisierung an sich, ist mit verstärkten und zusätzlichen negativen Auswirkungen auf Individuen und Gesellschaft zu rechnen.

In der vorliegenden Arbeit werden anhand aktueller wissenschaftlicher Publikationen die Funktionsweise und Fähigkeiten eines wahrscheinlichen automatisierten Systems entworfen. Dieses wird auf gesellschaftliche Probleme hin untersucht und Lösungsansätze werden diskutiert.

Identifizierte Probleme sind eine inhärent gegen das Prinzip der Datensparsamkeit verstoßende, umfangreiche Erhebung von personenbeziehbaren Daten, eine gesteigerte Informationsasymmetrie zwischen Betroffenen und Überwachenden, ein daraus resultierender massiver Eingriff in das Recht auf informationelle Selbstbestimmung sowie eine Steigerung der selbstdisziplinierenden Wirkung auf Betroffene. Die Arbeit kommt zu dem Schluss das bestehende Ansätze zur Anonymisierung von Daten technisch und konzeptuell ungenügend sind.

Mit der Technik der Verhaltenserkennung und -bewertung geht außerdem die Gefahr einer institutionalisierten Diskriminierung Betroffener

einher, die – wenn überhaupt identifiziert – nur schwer oder gar nicht zu unterbinden ist.

Entgegen der allgemeinen Argumentation, kann der Einsatz von OperateurInnen die umfangreiche Automatisierung nicht legitimieren und eine Verhinderung automatisierter Entscheidungen zum Nachteil Betroffener nicht sicherstellen. Zum einen können für eine mündige Entscheidung konzept- und technikbedingt keine adäquaten Informationen bereitgestellt werden. Zum anderen muss gegenüber der Assistenz durch das System aus psychologischen Gründen mit einem übersteigerten Vertrauen gerechnet werden, das zu einer mangelnden Überprüfung der – ohnehin unzulänglich überprüfbaren – Darstellung und Entscheidungen des Systems führt. Das Vertrauen kann nur teilweise und mit hohem Trainingsaufwand angepasst werden.

Nicht nur Auswirkungen einzelner Maßnahmen, sondern das Stattfinden automatisierter Überwachung allgemein, können eine positive gesellschaftliche Entwicklung gefährden. Die Technik birgt außerdem die Gefahr als Werkzeug zur Unterdrückung missbraucht zu werden.

Inhaltsverzeichnis

1 Einleitung

Ermöglicht durch die Entwicklung zahlreicher Techniken wird sowohl staatliche als auch private Überwachung immer umfassender, ungerichteter und tiefgreifender. Sowohl der traditionelle öffentliche Raum als auch der digitale Raum werden aus Sicherheitsgründen, aber zunehmend auch aus wirtschaftlichen Interessen, beobachtet und Geschehnisse ausgewertet.

Das Konzept „Überwachung" ist weder generell zu befürworten, noch grundsätzlich abzulehnen. Überwachung ist oft etwas Positives und Notwendiges. Überwachung von Prozessen in der Natur kann Katastrophen verhindern, Überwachung von Prozessen in der Industrie kann Produktqualität sichern und in Krankenhäusern kann Überwachung von medizinischen Werten schnelles Handeln ermöglichen.

Ob Überwachung ethisch ist, hängt davon ab, wer die Überwachung zu welchem Zweck durchführt, ob sie *verhältnismäßig* ist und ob für den Zweck die Notwendigkeit genau dieser Art und Ausprägung von Maßnahme besteht. Nach dem Prinzip der Verhältnismäßigkeit sind die Effektivität der Maßnahme und die resultierenden Einschränkungen der Betroffenen gegeneinander abzuwägen.[1]

Das Prinzip der Verhältnismäßigkeit ist in Bezug auf Videoüberwachung auch im Bundesdatenschutzgesetz wiederzufinden. Vor der Installation einer Maßnahme müssen sowohl Bedarf und Effektivität, als auch mögliche Einschränkungen der Grundrechte evaluiert werden. Eine Effektivität des Konzeptes Videoüberwachung im öffentlichen Raum konnte bisher jedoch nicht belegt werden.[2] In der Praxis erbrachte Videoüberwachung im öffentlichen Raum weder den erhofften präventiven Nutzen,

1 **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 7.

2 Vgl. z. B. **Rothmann**, „Zur Evaluation der Sicherheitstechnischen Eignung von Videoüberwachung. Regionale Defizite, internationale Standards, methodische Herausforderungen".

noch half sie signifikant[3] bei der Aufklärung von Straftaten. Entscheidend für dieses Ausbleiben waren das Ignorieren der Kameras durch GewalttäterInnen und bewusstem Einplanen dieser bei Sachbeschädigung und anderer Kriminalität. Empirisch nachgewiesen wurden jedoch negative Effekte auf Individuen und Gesellschaft.[4] Weder die Effektivitätsmängel[5], noch die negativen Auswirkungen[6] spiegeln sich jedoch bisher im öffentlichen Diskurs wider.

Obwohl Diskussionen um Nutzen sowie ethische, datenschutzrechtliche und gesellschaftliche Probleme von Videoüberwachung noch nicht abgeschlossen sind, fand nahezu weltweit ein massiver Ausbau von Videoüberwachung statt.

Dass dieser weiter vorangetrieben wird und darüber hinaus an Mitteln geforscht wird, Videoüberwachung doch noch zu einem effektiven Instrument zu machen, zeigt, dass ohne objektive Grundlage ein starker Glaube an das Konzept „Videoüberwachung" vorherrscht.

Nachdem Videoüberwachung digitalisiert und computerisiert wurde, so dass Algorithmen zum Einsatz kommen können, wird intensiv daran geforscht, sie zu automatisieren. Man verspricht sich von Automatisierung Effektivitätsprobleme zu lösen, dem „gesteigerten Sicherheitsbedürfnis" nachzukommen[7] und nicht zu Letzt, Videoüberwachung zu einem ökonomisch effizienten Mittel zu machen, mit dem Arbeitskraft eingespart werden kann. Es wird außerdem erwartet, technische Möglichkeiten aus-

3 Vgl. z. B. **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy", S. 323.

4 **Raab**, „Impact of Surveillance on civil liberties and fundamental rights", S. 267.

5 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 94.

6 **Möllers** ; **Hälterlein**, „Privacy issues in public discourse: the case of "smart" CCTV in Germany", S. 10.

7 Oftmals werden auch in wissenschaftlichen Veröffentlichungen zu Techniken der Automatisierung von Videoüberwachung die terroristische Anschläge in den USA im September 2001 als Ursache für ein gesteigerte Sicherheitsbedürfnis angeführt. Vgl. z. B. **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems", S. 182.

nutzen zu können, um den Datenschutzbestimmungen besser nachkommen zu können.

Neben der Forschung an einzelnen Techniken existieren in Europa zahlreiche von der EU oder den Mitgliedsstaaten geförderte Forschungsprojekte zur Automatisierung von Videoüberwachung wie zum Beispiel *INDECT*, *ADABTS*, *CamInSens*, *ADIS*, *APFeL*, *ASEV* oder *SINOVE*.[8] *INDECT* (Intelligent information system supporting observation, searching and detection for security of citizens in urban environment) hat in dieser Aufzählung eine besondere Bedeutung, da dessen Ziel nicht nur die Automatisierung von Videoüberwachung ist, sondern auch, andere Überwachungsinfrastrukturen in einem umfangreichen Überwachungswerkzeug zu integrieren.

Die Automatisierung mittels Algorithmisierung verändert grundsätzlich den Charakter von Videoüberwachung. Die computerisierte Automatisierung ermöglicht nicht nur eine qualitative und quantitative Steigerung der Erhebbarkeit und Verarbeitbarkeit von Informationen, sondern birgt gegenüber Überwachung durch Menschen eine Reihe weiterer möglicher bisher nicht bedachter Effekte, die sich sowohl aus den Charakteristika der Techniken und ihrem Zusammenwirken, als auch in der Wirkung des Wissens über Automatisierung auf Betroffene ergeben können.

Solche Effekte müssen identifiziert werden, um eine Bewertung der Verhältnismäßigkeit zukünftiger Überwachungsmaßnahmen vornehmen zu können – viel mehr noch aber, um rechtzeitig Einfluss auf technische Entwicklungen nehmen zu können und EntwicklerInnen für die Auswirkungen ihrer Arbeit zu sensibilisieren.

Dies zu tun, liegt in besonderem Maße auch in der Verantwortung der Informatik als Wissenschaft, da sie grundsätzlich die Mittel zur Automati-

8 Die hier beispielhaft genannten Projekte werden bis auf *INDECT* alle vom Bundesministerium für Bildung und Forschung gefördert. Vgl. URL: `http://www.bmbf.de/de/14395.php` (14. März 2015).

sierung bereitstellt. Allein sie ist mit dem Verständnis über die Techniken in der Lage, die Technik kritisch einzuschätzen und mögliche Probleme zu identifizieren.

Forschungsprojekte wie *INDECT* kommen dieser Verantwortung nicht nach. Ethische, gesellschaftliche und rechtliche Überlegungen werden nicht über einen späteren Einsatz der Techniken angestellt, sondern beziehen sich ausschließlich auf die Arbeit im Rahmen des Forschungsprojektes. Für die genaue Ausgestaltung und den rechtmäßigen Einsatz seien die Benutzer verantwortlich.[9]

Datenschützer wie der Berliner Datenschutzbeauftragte Alexander Dix äußern, dass derartige Projekte „im Grunde [drohen] Geld zu verschwenden, wenn die Ergebnisse hinterher nicht rechtskonform angewendet werden können".[10] Dass neben Projekten wie *INDECT* trotzdem eine Vielzahl ähnlicher Projekte staatlich finanziert wird, lässt erwarten, dass weiterhin Geld „verschwendet" wird, derartige Techniken nicht rechtmäßig eingesetzt werden oder aber Überwachungstechniken an Länder exportiert werden, in denen der Einsatz trotz der zu erwartenden Auswirkungen durchgesetzt werden kann.

1.1 Zielsetzung und Methode

Ziel der vorliegenden wissenschaftlichen Arbeit ist es, Probleme automatisierter Videoüberwachung im öffentlichen Raum zu identifizieren, die bei der Entwicklung, dem Export und der Abwägung der Verhältnismäßigkeit von Überwachungsmaßnahmen beachtet werden müssen. Im Zentrum der Betrachtung stehen Probleme, die sich aus den zugrundeliegenden Techniken, ihrem Zusammenspiel oder im weiteren Sinne aus der

9 Europäisches Parlament: *Parlamentarische Anfrage E-3190/2010 – Antwort von Herrn Tajani im Namen der Kommission* (2010) `http://www.europarl.europa.eu/sides/getAllAnswers.do?reference=E-2010-3190&language=DE`.

10 **3Sat (Hrsg.):** *Kulturzeit : INDECT*.

Automation an sich ergeben und einen negativen Effekt auf betroffene Individuen oder die Gesellschaft als Ganzes haben. Für die identifizierten Probleme werden außerdem Lösungsansätze diskutiert.

Die Vorgehensweise zur Identifizierung der Probleme unterliegt der gleichen Schwierigkeit, der die Technikfolgeabschätzung generell unterliegt: Der Spannung zwischen „realen Gestaltungsmöglichkeiten" und der „Verfügbarkeit verlässlichen Folgenwissens" – auch *Collingridge-Dilemma* genannt.[11] Zur Folgeabschätzung in späten Phasen der Entwicklung stehen zwar genügend Informationen bereit, Erkenntnisse können jedoch nur schwer berücksichtigt werden, da die Folgen bereits eingetreten sind oder eine Veränderung z. B. aus Interessenkonflikten nicht mehr realisierbar ist. Führt man die Folgeabschätzung in früheren Phasen durch, kann sie zwar Einfluss nehmen, stützt ihr Wissen jedoch auf unzureichende Informationen oder Spekulationen. In der Praxis besteht zwischen diesen beiden Extremen jedoch ein „fließender Übergang", so dass ein konstruktiver Umgang mit dem Problem möglich ist.[12]

Die Entwicklung automatisierter Videoüberwachung befindet sich noch in einer frühen Phase. Grundlagen für die Herangehensweise haben sich jedoch schon so weit ausdifferenziert, dass die generelle Funktionsweise der Techniken für einzelne Aspekte der Überwachungsaufgabe und auch deren Zusammenspiel in zukünftigen Systemen mit dem grundsätzlichen Verständnis für informationstechnische Systeme bereits abzuschätzen ist. Im Folgenden wird eine solche Abschätzung vorgenommen und deren Ergebnis nach verschiedenen Kriterien untersucht. Dabei wird Videoüberwachung und ihre Automatisierung vor allem im deutschen und europäischen Kontext betrachtet und daher Begrifflichkeiten und Argumentation meist an deutsches oder europäisches Recht angelehnt. Die abschließen-

11 **Grunwald**, *Technikfolgenabschätzung: eine Einführung*, S. 165.
12 Ebd., S. 166.

den Betrachtungen beschränken sich jedoch nicht auf den europäischen Raum.

Um die Abschätzung eines möglichen Überwachungssystems möglichst plausibel durchzuführen, wurde der Diplomarbeit eine intensive Recherche wissenschaftlicher Veröffentlichungen über die Automatisierung verschiedenster Aspekte der Überwachungsaufgabe vorangestellt. Anhand von Medienberichten der letzten acht Jahre, die hauptsächlich in Zusammenhang mit dem *INDECT*-Projekt standen, wurde eine Reihe von Begriffen herausgearbeitet, die viele Aspekte der Automatisierung von Videoüberwachung abdecken. Anhand dieser Begriffe wurde nach wissenschaftlichen Publikationen gesucht. Viele der gefundenen Veröffentlichungen – besonders aus dem Gebiet der Computervision – beziehen sich bereits direkt auf Videoüberwachung und referenzierten weitere Quellen, so dass sich ein umfangreiches Bild des Forschungsstandes, zumindest der publizierten Forschung, ergab. Die angestrebte Automatisierung umfassen jedoch nicht nur das Bildverstehen. Es zeichneten sich auch weitere Trends wie Vernetzung, Dezentralisierung und das Integrieren weiterer Informationsquellen ab. Dies zeigte die Notwendigkeit auf, nicht nur das visuelle System eines denkbaren Überwachungssystems zu berücksichtigen, sondern auch weitere Aspekte in die Abschätzung einfließen zu lassen. Da eine Bewertung aus rein technischer Sicht für die Identifizierung gesellschaftlicher Probleme nicht hinreichend ist, wurde sich mit grundlegenden soziologischen, humangeografischen, juristischen und psychologischen Perspektiven auf Videoüberwachung beschäftigt. Auf diese Weise wurden Auswirkungen auf Individuen und Gesellschaft sowie Eigenschaften des Konzeptes Videoüberwachung herausgearbeitet. Besonders zuträglich war eine Veröffentlichung von Francisco Klauser, in der die Struktur der Videoüberwachung und deren Wirkung auf den überwach-

ten Raum beschrieben wird.[13] Impulsgebend waren außerdem Erkenntnisse der Ingenieurpsychologie aus dem Bereich der Mensch-Technik-Interaktion. Auf Grund dieser musste vermutet werden, dass sowohl die Darstellung der Analyseergebnisse als auch die Assistenz durch Automatisierung einen Effekt auf die OperateurInnen und deren Umgang mit dem System hat.
Das zuvor konstruierte Bild automatisierter Videoüberwachung konnte dann auf die Effekte hin untersucht werden und mit informatischem Wissen bewertet werden. Dazu wurden Veröffentlichungen herangezogen die diese Effekte im Einsatz von solchen Systemen nachweisen, die mit einem automatisierten Videoüberwachungssystem vergleichbar sind. Übertragbar waren Ergebnisse von Jennifer Bahner[14] zu übersteigertem Vertrauen in Automation sowie von Lucas Introna und David Wood[15] zu diskriminierenden Aspekten von Gesichtserkennungsalgorithmen.
Für die Verständlichkeit erschien es sinnvoll, die Arbeit nicht chronologisch nach der eben beschriebenen Vorgehensweise zu strukturieren, sondern wie im Folgenden beschrieben, den Hauptteil in vier Kapitel zu gliedern.

1.2 Struktur des Buches

In Kapitel 2 „Vorbetrachtungen zu Videoüberwachung“ wird in das Themenfeld Videoüberwachung eingeführt. Um das Ausmaß und die Relevanz zu veranschaulichen, wird ein kurzer geschichtlicher Abriss über die Ausbreitung von Videoüberwachung gegeben. Außerdem wird untersucht, inwiefern die Effektivität berücksichtigt wurde und das Problem-

13 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*.
14 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias“.
15 **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems“.

bewustsein in der Bevölkerung vorhanden ist. Anschließend werden die Begriffe *manuelle* und *automatisierte Videoüberwachung* anhand der technischen Entwicklungsgeschichte eingeführt, ohne dabei schon auf technische Details einzugehen. Nachdem der Begriff „öffentlicher Raum" eingeführt wird, in dem die hier betrachtete Videoüberwachung zum Einsatz kommen soll, werden das Konzept und die Probleme *manueller* Videoüberwachung dargelegt. Dazu wird die Aufgabe der OperateurInnen beschrieben, die Bedeutung des Datenschutzes und der informationellen Selbstbestimmung für Videoüberwachung dargestellt und die Auswirkung auf Individuen und Gesellschaft aufgezeigt.
In Kapitel 3 „*Techniken und Konzepte automatisierter Videoüberwachung*" werden dann Techniken und Konzepte vorgestellt, die der Automatisierung der Videoüberwachung dienen und Teil eines Überwachungssystems sein können. Die Techniken zusammen in einem Kapitel und nicht jeweils im Zusammenhang der resultierenden Probleme vorzustellen, ist sinnvoll, da viele Technikaspekte für mehrere Probleme relevant sind und sich einige Probleme erst aus dem Gesamtbild erschließen. Um eine klare Struktur zu erzielen, werden bei den Beschreibungen noch keine Bewertungen vorgenommen, wenn diese nicht bereits in den Veröffentlichungen selbst explizit angesprochen werden. Nachdem einige Begriffe für Automatisierungstechniken diskutiert werden, wird zunächst die Verhaltenserkennung vorgestellt. Ihr hierarchischer Aufbau und die verschiedenen Abstraktionsebenen mit Techniken der Bildverarbeitung und der künstlichen Intelligenz werden skizziert. Ein besonderer Fokus wird auf die Gewinnung und die Form von Modellen gelegt, die Verhalten oder Vergleichbares repräsentieren und zur Erkennung benutzt werden sollen. Anschließend werden weitere Techniken vorgestellt, mit denen Informationen aus den Bildern bzw. dem überwachten Raum extrahiert werden sollen. Dabei wird auch auf Identifizierung anhand biometri-

scher Merkmale eingegangen. Danach werden Kameras, deren Netzwerkfähigkeit und Rechenkapazität und weitere Sensoren vorgestellt. Im Anschluss wird die Möglichkeit einer zeitlich rückwärts gerichteten Überwachung beschrieben, die durch Speicherung der Daten ermöglicht wird. Die Möglichkeiten der Einbeziehung und Analyse weiterer Datenquellen, die nicht im Rahmen der Überwachung selbst erhoben wurden, kann nur kurz angeschnitten werden. Dem folgend werden Darstellungsansätze der Analyseergebnisse und Kamerabilder für die OperateurInnen beschrieben. In einem eigenen Abschnitt werden Techniken behandelt, die dem Datenschutz dienen sollen.

Da sich der Charakter automatisierter Videoüberwachung erst im Zusammenspiel der Techniken zeigt, wird in Kapitel 4 basierend auf den vorgestellten Techniken und Veröffentlichungen über Systemarchitektur, ein technisch wahrscheinliches Gesamtsystem Ω entworfen. Dieses wird mit dem vollen technischen Potential entworfen, ohne dass auf rechtliche Einschränkungen Rücksicht genommen wird. Dies erscheint vor dem Hintergrund aktuell zu beobachtender umfangreicher, tiefgreifender und ungerichteter Überwachungspraxis durch öffentliche und private Stellen, aber vor allem Geheimdienste nicht nur legitim, sondern für eine umfangreiche und realistische Folgeabschätzung zwingend erforderlich.

Anhand der Struktur und Funktionsweise von Ω werden im Kapitel 5 die Problemfelder hergeleitet und beschrieben. Obwohl die Probleme zum Teil stark miteinander verwoben sind und sich gegenseitig bedingen, werden die einzelnen Probleme in eine möglichst logische Reihenfolge gebracht. Dazu wurde eine Gliederung in vier Kapitel gewählt:

Zunächst werden Folgen für Datenschutz und informationelle Selbstbestimmung hergeleitet, die sich unmittelbar aus der Gestaltung des Systems ergeben. Es wird analysiert, inwieweit durch die Automatisierung neben Raumüberwachung auch Überwachung von Individuen ermög-

licht wird. Es wird auf die datenschutzrechtlich relevante Verkettbarkeit und Zweckgebundenheit der erhobenen Daten eingegangen, die Einsichtnahme und Überprüfbarkeit durch Betroffene besprochen und kurz die Datensicherheit angeschnitten. Gesondert untersucht wird, inwiefern sich aus der Funktionsweise, den Zielen und dem Überwachungskonzept ein gesteigerter Datenbedarf ergibt, der in Spannung mit dem Datenschutzgrundsatz der Datensparsamkeit steht. Anschließend werden die zuvor beschriebenen technischen Maßnahmen zum Datenschutz diskutiert.
Im zweiten Abschnitt des Kapitels wird die Rolle der OperateurInnen als Legitimation einer umfangreichen Automatisierung analysiert. Es wird erörtert, inwieweit in Hinblick auf die filternde und interpretierende Funktionsweise von Ω, von einer mündigen Entscheidung ausgegangen werden kann. Außerdem wird die Übertragbarkeit des empirisch nachgewiesenen psychologischen Effekts eines übersteigerten Vertrauens (*complacency*) in Assistenzsysteme auf ein mögliches dem System Ω entgegengebrachtes Vertrauen hergeleitet.
Im dritten Abschnitt wird untersucht, inwiefern Algorithmen entgegen der allgemeinen Annahme nicht als objektiv und wertfrei angesehen werden können. Ausgehend von dem Beispiel diskriminierender Gesichtserkennung werden anhand der Funktionsweise von Ω Einfallstore für Diskriminierung erörtert und Möglichkeiten zur Verhinderung diskutiert.
Abschließend werden individuelle und gesellschaftliche Auswirkungen diskutiert, die sich nicht direkt aus der Funktionsweise, sondern aus den möglichen Fähigkeiten und den Erwartungen der Betroffenen ergeben. Es wird erörtert, inwiefern durch den Einsatz von automatisierter Videoüberwachung, die Auswirkungen, die schon bei *manueller Videoüberwachung* auftreten, verstärkt werden. Es wird außerdem untersucht inwiefern die Steigerung der Effekte selbst dann zu erwarten ist, wenn im konkreten Fall *manuelle* oder wenig automatisierte Videoüberwachung

stattfindet. Abschließend wird diskutiert, welche Maßnahmen zur Vermeidung der Probleme durch Überwachende getroffen werden könnten.

2 Vorbetrachtungen zu Videoüberwachung

2.1 Ausbreitung und Effektivität

Ausbreitung

1965 wurden zum ersten mal Videokameras in Großbritannien als Assistenz zur Steuerung von Ampeln durch nur eine Person eingesetzt.[16] Die Ausbreitung beschleunigte sich ab den späten 1970er Jahren[17] und wurde nach den „London bombings" im Jahre 1990 immer stärker vorangetrieben.[18] Als Reaktion auf terroristische Anschlägen auf New York im Jahr 2001 begann ein massiver Ausbau staatlicher und privater Videoüberwachung im öffentlichen Raum.[19] Der stärkste Zuwachs von Videoüberwachung war in den vergangenen 20 Jahren im Vereinigten Königreich Großbritannien und Nordirland zu beobachten. In London waren 2004 40% des öffentlichen Raumes videoüberwacht. Nach neuesten Schätzungen der British Security Industry Association (BSIA) im Juli 2013 existieren im Vereinigten Königreich ca. 4 - 5,9 Millionen Überwachungskameras.[20] Der Kartenausschnitt von London in Abb. 1 aus dem letzten Jahr, in den lediglich die während einer siebenstündigen Begehung gesichteten Kameras eingezeichnet wurden, veranschaulicht das Ausmaß.

Auch im übrigen Europa fand ein immenser Ausbau statt. 2002 ergaben Stichproben von mehr als 1.400 öffentlich zugänglichen Orten (wie Läden, Bahnhöfe, Kinos, Banken, etc.) in den Hauptstädten von sechs europäischen Staaten, dass 29% aller Gebäude und Einrichtungen ein Videoüberwachungssystem benutzten.[21] In Paris hat sich die Anzahl der

16 **Kreissl**, „The effectiveness of surveillance in preventing and detecting crime and terrorism", S. 178.
17 **Hempel** ; **Töpfer**, „Videoüberwachung in Europa : Abschlussbericht", S. 4.
18 **Senior**, *Protecting Privacy in Video Surveillance*, S. VII.
19 **Deutscher Bundestag (Hrsg.):** *Drucksache 17/2750*, S. 3.
20 **Reeve**, *BSIA attempts to clarify question of how many CCTV cameras there are in the UK*.
21 **Hempel** ; **Töpfer**, „Videoüberwachung in Europa : Abschlussbericht", S. 4.

Kameras innerhalb des Jahres 2009 vervierfacht.[22] In Deutschland sind aktuell allein im bahnpolizeilichen Aufgabenbereich der Bundespolizei 3.000 Kameras auf rund 300 Bahnhöfen installiert.[23] In Bayerns Kommunen existieren nach neuesten Berichten 17.000 Überwachungskameras.[24] Der Ausbau wird nicht nur in Europa vorangetrieben. Videoüberwachung prägt inzwischen nahezu weltweit die Stadtbilder – besonders in Metropolen. In New York wurde zu Beginn des Jahrtausends innerhalb von zwei Jahren die Anzahl der „crime fighting cameras" von 250 auf 3.000 in einem Netzwerk integrierter Kameras aufgestockt.[25] In Chicago wurde zwischen November 2008 und Februar 2009 ein eigenes Glasfasernetzwerk zur Videoüberwachung aufgebaut.

Der Einsatz von Videoüberwachung nimmt auch außerhalb der westlichen Welt zu. Berichten zu Folge sind beispielsweise in China bereits mehr als 2,7 Millionen Kameras montiert, davon allein in Peking fast 300.000 Geräte.[26]

Evaluierung der Effektivität

Vor und während der massiven Ausbreitung fand kaum Evaluierung des Konzepts „Videoüberwachung" zur *Prävention* und *Strafverfolgung* statt[27] und auch die Zweckmäßigkeit von Maßnahmen im konkreten Fall wurden vor ihrer Installation selten adäquat evaluiert.[28] Die Annahme des *präventiven Nutzens*, dass Menschen keine Verbrechen begehen würden, oder sie anderswo begingen[29], weil sie fürchten, von aktiver Überwa-

22 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy".
23 **Deutscher Bundestag (Hrsg.):** *Drucksache 17/2750*, S. 4.
24 **Kannenberg**, *Bayerische Datenschützer: Schon 17.000 kommunale Überwachungskameras*.
25 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy", S. 323.
26 **Erling**, *Chinas Polizei will den totalen Überblick*.
27 **Kreissl**, „The effectiveness of surveillance in preventing and detecting crime and terrorism", S. 179.
28 **Raab**, „Impact of Surveillance on civil liberties and fundamental rights", S. 307.
29 **Senior** et al., „Enabling Video Privacy through Computer Vision", S. 50.

Abb. 1: Unvollständige Erfassung von Kameras (·) und überwachter öffentlicher Raum (●) in London südlich der Themse im Mai 2012.

chung gefasst oder nachträglich auf dem Videoaufnahmen identifiziert zu werden, wurde mit Studien nie eindeutig belegt. Es konnten keine konsistenten Beweise zur Reduzierung von Kriminalität erbracht werden.[30]
Eine britische Metastudie über 22 britische und US-Amerikanische Studien zur Effektivität von Videoüberwachung ergab, dass der *präventive* Einsatz die Kriminalität durchschnittlich lediglich um 4 Prozent verringert hätte.[31] Die Aussagekraft von Studien wird jedoch generell in Frage gestellt, da diese unter unzureichender Beachtung der Rahmenbedingungen durchgeführt werden, die die Effektivität beeinflussen.[32]
Auch als Unterstützung zur *Strafverfolgung* zeigte Videoüberwachung in der inzwischen langjährigen Praxis eine geringe Effektivität. Es werden zwar Fälle zitiert, in denen Videobilder unzweifelhaft dazu beigetragen haben, verdächtige Personen – meist für schwerwiegende Straftaten – zu

30 **Kreissl**, „The effectiveness of surveillance in preventing and detecting crime and terrorism"; **Rothmann**, „Zur Evaluation der Sicherheitstechnischen Eignung von Videoüberwachung. Regionale Defizite, internationale Standards, methodische Herausforderungen".

31 **Welsh** ; **Farrington**, *Crime Prevention Effects of Closed Circuit Television: A Systematic Review*.

32 **Apelt** ; **Möllers**, „Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung", S. 589.

verhaften, die vermeintlich für die Effektivität sprechen, der New Yorker Bürgermeister Bloomberg gab jedoch an, dass 2008 mit Hilfe von Videoüberwachung nur 3 Prozent der Straßenraube und insgesamt nur 1.000 Straftaten („crimes") gelöst werden konnten. Dies bedeutet im Jahr weniger als einen gelösten Fall pro 1.000 Kameras.[33] Die Evaluation der Videoüberwachung der Berliner Verkehrsbetriebe (BVG) ergab, dass im monatlichen Mittel von allen bekannten Vorfällen in nur 2,1% der Fälle die Aufnahmen zur Identifizierung führten.[34] Besonders für die Aufklärung von Sachbeschädigungen dienten die Aufnahmen nicht, da vorhandene Kameras von den TäterInnen eingeplant werden.

Rechtmäßigkeit

Nach § 6b des deutschen Bundesdatenschutzgesetzes ist die Beobachtung öffentlich zugänglicher Räume mit Videoüberwachung nur zulässig, soweit sie für einen konkret festgelegten Zweck erforderlich ist und keine Anhaltspunkte bestehen, dass schutzwürdige Interessen der Betroffenen überwiegen. Vor und zur Verlängerung der Laufzeit einer Überwachungsanlage muss demnach das Prinzip der Verhältnismäßigkeit beachtet werden. Der Notwendigkeit und Tauglichkeit steht die Schwere des Grundrechtseingriffes gegenüber.

In Deutschland findet Videoüberwachung jedoch oft statt, ohne dass dabei auf diese engen Grenzen des Datenschutzes geachtet wird.[35] Hier verstießen von 2008 bis 2010 99% der 3345 überprüften von Kommunen betriebenen Kameras massiv gegen datenschutzrechtliche Vorschriften. Auch nichtöffentliche Betreiber von Videoüberwachung verstoßen „seit vielen Jahren [...] massiv gegen datenschutzrechtliche Vorschriften".[36] In

33 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy", S. 323.
34 **Hempel** ; **Alisch**, *Evaluation der 24-Stunden-Aufzeichnung in Ubahnstationen der Berliner Verkehrsbetriebe (BVG) : Zwischenbericht*.
35 **Meyer**, *Zweifelhafter Notanker : Videoüberwachung in Schulen*.
36 **Wahlbrink**, *Zahlreiche Rechtsverstöße bei der Videoüberwachung : Wahlbrink: Behörden und Kommunen ignorieren Datenschutzgesetz*.

London, so wird geschätzt, seien nahezu 80% der privat betriebenen Kameras nicht rechtskonform.[37]

Mangelndes Problembewusstsein und Akzeptanz

Videoüberwachung „scheint heute der Königsweg zur Bewahrung öffentlicher Ruhe und Ordnung" geworden zu sein, äußerte Nogala schon im Jahre 2000.[38] Nach wie vor werden regelmäßig politische Forderungen nach Videoüberwachung laut und von geplanten staatlichen und privaten Maßnahmen berichtet. Jüngst gab die Deutsche Bahn an, zusätzlich zu stationärer Videoüberwachung auch mit Kameras ausgestattete Flugdrohnen zum Einsatz bringen zu wollen.[39] Das Videoüberwachung weiterhin als Lösung propagiert wird mag daran liegen, dass Ergebnisse der Studien über ausbleibende Effektivität und die mangelnde Rechtmäßigkeit medial kaum präsent sind. Klauser stellt fest, dass die Erkenntnisse kaum Erwähnung im Rahmen der Diskussion zur Videoüberwachung finden.[40]
Auch in der Bevölkerung scheint die Akzeptanz von Videoüberwachung stark vorhanden zu sein, jedoch ohne eine nachweisbare Effektivität und sogar ohne dass Betroffene tatsächlich von einem gesteigerten Sicherheitsgefühl berichten würden.[41] Doch diese bis jetzt ausgebliebene „Rebellion der Betroffenen" gegen ihre „Durchleuchtung und Überwachung sollte" nach Tichy „nicht als Akzeptanz oder gar als Zustimmung interpretiert werden" sondern eher als „mangelndes Problembewußtsein [sic]" angesehen werden.[42]

---

37 **Senior**, *Protecting Privacy in Video Surveillance*, S. 39.
38 **Nogala**, *Der Frosch im heißen Wasser : Die Trivialisierung von Überwachung in der informatisierten Gesellschaft des 21. Jahrhunderts*.
39 **Futurezone.at (Hrsg.):** *Deutsche Bahn will Drohnen gegen Sprayer*.
40 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 94.
41 **Apelt** ; **Möllers**, „Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung", S. 587.
42 **Tichy** ; **Peissl**, *Beeinträchtigung der Privatsphäre in der Informationsgesellschaft*, S. 13.

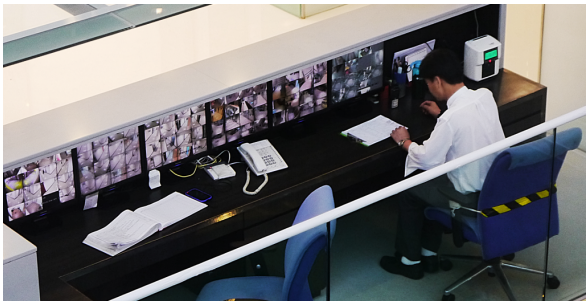


Abb. 2: Manuelle Videoüberwachung.

2.2 Verlauf der technischen Entwicklung

In der Literatur wird die Entwicklungsgeschichte von Videoüberwachungssystemen oft nach dem Grad der Digitalisierung strukturiert. Helmut Schwabach et al.[43] und David d'Angelo et al.[44] unterscheiden nach diesem Kriterium drei Generationen. Überwachungssysteme der *ersten Generation* bestanden aus analogem Equipment. Die Kameras lieferten meist verschwommene, grobkörnige, oft viel zu dunkle Bilder. Das Videosignal wurde analog zu einem zentralen „back-end system" übertragen und dort den OperateurInnen wie in Abb. 2 auf Monitoren angezeigt oder für die spätere Sichtung archiviert. Meist geschah dies auf Magnetbändern, deren praktische Nachteile wohl keiner Erwähnung mehr bedürfen.
Systeme der *zweiten Generation* sind teildigitalisiert. Sie benutzten digitale Komponenten zur Verarbeitung und Analyse der Videodaten in Echtzeit mit automatisierter Detektion von Ereignissen und der Generierung von Alarmen.

43 **Schwabach** et al., „Distributed Embedded Smart Cameras for Surveillance Applications".

44 **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces".

Die *dritte Generation* wird durch die vollständige Digitalisierung des Systems charakterisiert. Digitale Kameras übermitteln gegebenenfalls bereits komprimierte Bilddaten über ein Computernetzwerk. Videodaten werden an intelligenten „hubs" gesammelt, verarbeitet und an die OperateurInnen oder ein Archiv weitergeleitet.[45]

Da Videoüberwachungssysteme im Folgenden nicht nur aus technischer, sondern auch aus gesellschaftlicher Perspektive betrachtet werden sollen, ist die Gliederung der Entwicklung anhand des rein technischen Aspekts des *Digitalisierungsgrades* nicht hinreichend.

Bei der im Folgenden entworfenen angemesseneren Unterteilung wird auch der *Grad der Automatisierung* mit einbezogen. Mit diesem zusätzlichen Kriterium verschiebt sich der Fokus von der Art der technischen Umsetzung auf die Struktur der Aufgabe und die Rolle der OperateurInnen. Zur Unterscheidung sollen römischen Zahlen der Bezeichnung der drei Generationen dienen. Die Generationen I und II werden im Präteritum beschrieben, da sie als Entwicklungsstände der Vergangenheit angesehen werden können. Eigenschaften der Generation III haben sich noch nicht ausdifferenziert und werden daher im Präsens beschrieben. Genau wie zwischen den Generationen *zwei* und *drei* von Schwabach und d'Angelo ist auch der Übergang von Generation II zu III fließend. Die Unterteilung dient daher eher der Veranschaulichung der Entwicklungstendenzen und nicht der Kategorisierung real existierender Systeme.

Systeme der *Generation I* waren *technische Voraussetzung* zur ersten visuellen Überwachung aus der Ferne. Sie dienten lediglich der Aufnahme und Bereitstellung der Videobilder. Bei ihrem Einsatz erfüllten menschliche OperateurInnen die damals neuartige Überwachungsaufgabe vollständig und selbstständig. Sie bestand aus *Auswahl* der näher zu betrachtenden Bilder, ihrer *Interpretation* sowie Abwägung und Einleitung entsprechen-

45 **Schwabach** et al., „Distributed Embedded Smart Cameras for Surveillance Applications", S. 69.

Abb. 3: Das Betreten der Gleise wird über simple Bewegungdetektion registriert.

der *Maßnahmen*. Eine Digitalisierung der Systems war hierfür nicht nötig.

Systeme der *Generation II* waren in der Lage den OperateurInnen bei ihrer, ansonsten gegenüber der Generation I unveränderten Aufgabe, zu *unterstützen*. Sie konnten OperateurInnen nach simplen Regeln alarmieren, die keiner komplexen semantischen Interpretation der beobachteten Geschehnisse bedurften (Abb. 3). Die Detektion beruhte dabei z. B. auf Bewegungserkennung, welche auch mit analogen Systemen umsetzbar war, oder auf einfachster Objektverfolgung, die in der Regel bereits eine computerisierte Verarbeitung digitalen Bildmaterials voraussetzte. Die Assistenz erfolgte ausschließlich bei der *Auswahl* der Kamerabilder, die eingehender betrachtet werden sollten, jedoch nicht bei der Interpretation der Geschehnisse.

Momentan wird der Übergang von Generation II zu Generation III vorangetrieben, deren Beschreibung und Auswertung Inhalt dieses Buches ist. Ziel der Systeme der *Generation III* ist es, den OperateurInnen mehr Aspekte der Überwachungsaufgabe *abzunehmen* und dem System zu übertragen bzw. diese ganz dem System zu überlassen. Man ist bestrebt, nicht nur die Auswahl der Bilder, sondern auch die Interpretation der Geschehnisse und die Auswahl der Maßnahmen zu automatisieren. Je nach Au-

tomatisierungsgrad ändert sich die Aufgabe der OperateurInnen qualitativ bzw. wird diese sogar weitestgehend aufgehoben. Bei umfangreicher Automatisierung besteht die Funktion der OperateurInnen dann darin, die Automation zu überprüfen. Dadurch ändert sich auch die Bedeutung der Rolle „OperateurIn", worauf in 5.2 näher eingegangen wird. Durch die Automatisierung ändert sich ferner auch die Quantität des Datenaufkommens und die Qualität der Datenauswertung bzw. der resultierenden Ergebnisse. Dieser Zusammenhang ist Gegenstand des Kapitels 5.1.
Systeme der *Generation III* können zusätzlich zu den vorherigen Aufgaben der OperateurInnen auch Aufgaben ausführen, die ein Menschen nicht oder zumindest nicht in einem solchen Umfang durchführen kann. Auf diese erweiterten Möglichkeiten wird in Kapitel 3 näher eingegangen. Videoüberwachung, die mit Systemen der Generation I und II durchgeführt wird, soll im Folgenden mit dem Retronym *manuelle Videoüberwachung* bezeichnet werden. Videoüberwachung mit Systemen, die eher der Generation III zuzuordnen sind, wird *automatisierte Videoüberwachung* genannt. Die Vielzahl der in den Medien und der Literatur gebrauchten Begriffe, die im weitesten Sinne Systeme der Generation III oder automatisierte bzw. computerisierte Videoüberwachung an sich bezeichnen, werden in Kapitel 3 diskutiert.

2.3 Konzept und Auswirkungen *manueller* Videoüberwachung

Vorbereitend für die Identifizierung von Problemen automatisierter Videoüberwachung wird im Folgenden *manuelle Videoüberwachung* charakterisiert. Dabei wird sich auf Aspekte konzentriert, die für die Auswertung automatisierter Videoüberwachung in Kapitel 5 von Bedeutung sind. Diese sind vor allem die Struktur der Überwachung, die maßgeblich die Aufgabe der OperateurInnen bestimmt und die Auswirkungen

auf Individuen und Gesellschaft. Dazu werden die Betrachtungen auf Videoüberwachung begrenzt, die im „öffentlichen Raum“ stattfindet. Dieser Begriff soll zunächst geklärt werden.

Der Begriff „öffentlicher Raum“

Das Verständnis des „öffentlichen Raums“ soll sich im Folgenden nicht durch Öffentlichkeit im eigentumsrechtlichen Sinne sondern durch die *Zugänglichkeit*, und vielmehr noch durch *soziale Begegnung*, die im Raum stattfindet, auszeichnen.[46] Klauser beschreibt öffentliche Räume in diesem Sinne als „zugängliche und nutzbare in Idealform gesellschaftlich geteilte Räume“.[47] Nach diesem Verständnis sind auch private, aber dennoch für die Allgemeinheit zugängliche Räume wie Kaufhäuser, Bahnhöfe oder private Spielplätze gemeint. Öffentliche Orte, an denen wenig oder kein gesellschaftliches Leben stattfindet (z. B. ein öffentliches Parkhaus) liegen eher außerhalb des Begriffsverständnisses. In Gesetzestexten wie dem Bundesdatenschutzgesetz (BDSG) findet der Begriff „öffentlich zugänglicher Raum“ Anwendung, der dem hier gemeinten Verständnis nahe kommt. Im Kapitel 5 ist daher mit *Videoüberwachung*, wenn nicht anders angegeben, die *Videoüberwachung im öffentlich zugänglichen Raum* gemeint. Soll die Öffentlichkeit besonders hervorgehoben werden, so wird – gleichbedeutend – verkürzend auch der Begriff „öffentlicher Raum“ gebraucht.

Eigenschaften von Videoüberwachung im öffentlichen Raum

Videoüberwachung im öffentlichen Raum ist nicht nur als personal- und kostensparendes technisches Instrument der Kriminalitätsbekämpfung zu verstehen.[48] Sie unterscheidet sich von Videoüberwachung im nichtöffentlichen Raum in Struktur, Ziel und – hier am relevantesten – in ih-

46 Vgl. **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 138.
47 Ebd., S. 164.
48 Ebd., S. 343.

rer Wirkung in den Raum und auf die Gesellschaft. Bei nichtöffentlicher Überwachungsmaßnahme z. B. zum Objektschutz, beobachten Kameras oft vom Rand des öffentlichen Bereichs einen Sichtpunkt (z. B. einen Hauseingang) oder eine Sichtachse (z. B. die „Außenhaut" eines Gebäudes).[49]

Videoüberwachung im öffentlichen Raum z. B. zur Innenstadtüberwachung, zur Überwachung von Massenveranstaltungen oder öffentlicher Verkehrsmitteln, verfolgt hingegen einen normativen Ansatz.[50] Häufig werden ganze Gebiete gesellschaftlichen Lebens meist mit mehreren frei schwenkbaren oder weitwinkligen Kameras mit sich überlappenden Bildbereichen beobachtet. Im Gegensatz zu anderen Formen der Überwachung (z. B. Telefonüberwachung) ist *manuelle Videoüberwachung* im öffentlichen Raum im Allgemeinen keine Überwachung von Einzelpersonen.[51] Der Fokus der Überwachung liegt auf dem Ort. Personen werden nur für den Zeitraum beobachtet, in dem sie sich im Bild befinden.[52] Dies heißt jedoch nicht, dass diese Form der Überwachung keinen Effekt auf Individuen hat.

2.3.1 Aufgabe der OperateurInnen

In Kapitel 2.2 wurde bereits die Veränderung der Aufgabe der OperateurInnen von Generationen I bis Generation III angesprochen. Die Aufgabe der OperateurInnen bei *manueller Videoüberwachung* soll im Folgenden genauer untersucht werden.

Bei Echtzeitüberwachung befinden sich OperateurInnen im Wesentlichen in einem Raum, der mit Bildschirmen ausgestattet ist, auf denen der be-

49 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 52.
50 Ebd., S. 71.
51 **Raab**, „Impact of Surveillance on civil liberties and fundamental rights", S. 255.
52 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 91.

obachtete Raum aus dem Blickwinkel der Kameras projiziert wird.[53] Die Kontrollaufgabe wird anhand vordefinierter Routinen erfüllt. Diese mehr oder weniger differenzierten Routinen sind nötig, damit OperateurInnen wissen, was überwacht werden soll und wann zu reagieren ist. Die OperateurInnen haben daher die Aufgabe, die dargestellten Bilder einzuschätzen und entsprechend dem Ziel der konkreten Videoüberwachungsmaßnahme zu reagieren bzw. weitere Maßnahmen einzuleiten.

Abstraktion des überwachten Raumes

Videoüberwachung kann öffentliche Räume in ihrer Komplexität nicht erfassen.[54] Es findet ein Transfer der Kontrolle vom eigentlichen auf einen abstrakten Raum statt.[55] Auf den Monitoren erscheint ein räumlich und informativ beschränkter Annäherungswert, anhand dessen die OperateurInnen die Situation einschätzen müssen. Bei diesem Prozess werden Informationen nach Klauser durch zwei Filter reduziert.[56] Dies ist zum Einen ein *technisch bedingter*, zum Anderen ein *mentaler Filter*.

Technischer Filter

Mit technischen Methoden wird versucht, den überwachten Raum zu modellieren.[57] Der Charakter dieses modellierten abstrakten Raumes wird geprägt durch eine ganze Reihe von „Mikroentscheidungen" und „Mikroaushandlungsprozessen" über die genauen Modalitäten, die während des gesamten Prozesses der Entwicklung und Installation einer Maßnah-

53 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 419.

54 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 122 ff.

55 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 420.

56 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 125 ff.

57 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 419.

me mit vielen Beteiligten stattfindet.[58] Die Vielzahl der Einflussfaktoren zeichnet sich schon ab, wenn man sich die Planung und Installation einer Maßnahme nur grob vorstellt: Basierend auf vorher getroffenen Annahmen über den zu überwachenden Raum werden Risikobereiche ausgewählt. Die Möglichkeiten der Kameraperspektiven werden z. B. von baulichen Gegebenheiten oder örtlichen Bestimmungen eingeschränkt. Darüber hinaus wird je nach Wahl der Kamera das Bild mit bestimmter Qualität aufgenommen und anschließend je nach Art, Anordnung und Größe der Bildschirme auf bestimmte Weise dargestellt. Die Reduzierung ist nicht nur menschlichen Entscheidungen geschuldet. Beispielsweise sind bestimmte sensorische Informationen schlichtweg nicht mit technischen Mitteln erfassbar. Derartige Informationen, die zur Kontextualisierung der Bilder essentiell sein könnten, stehen im abstrakten Raum nicht zur Verfügung.[59] Durch diese Reduzierung des komplexen öffentlichen Raumes auf flache Bilder findet a priori die erste Filterung und Dekontextualisierung statt.

Rekontextualisierung

Die überwachten Objekte oder Personen werden also aus ihrem Kontext herausgenommen und in den modellierten und somit vereinfachten Kontext zurückgesetzt.[60] Um den Übergang vom ursprünglichen Raum in den abgebildeten Raum verstehen zu können, müssen OperateurInnen diesen durch erneute Anreicherung mit Informationen *rekontextualisieren*. Diese Information entsprechen entweder standardisierten Abläufen, kommen von Sicherheitskräften vor Ort oder stammen aus der eigenen Erfahrung. Besonders die persönlichen Kenntnisse der OperateurInnen,

58 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 418.
59 **Norris** ; **Armstrong**, „CCTV and the social Structuring of Surveillance", S. 159.
60 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 425.

so wurde durch Auswertung von Erfahrungsberichten festgestellt, sind für eine angemessene Interpretation der Bilder von besonderer Bedeutung.[61]

Mentaler Filter

Zur technisch bedingten Reduzierung kommt noch ein „mentale Filter" hinzu. In der Praxis steht die Anzahl der auszuwertenden Bilder oftmals nicht im Verhältnis zur verfügbaren Anzahl der OperateurInnen bzw. zu ihren kognitiven Kapazitäten. In der Realität muss eine Person daher eine Vielzahl von Bildschirmen beobachten. Die Zuständigkeit einer Person, für über 50 Kameras und in Einzelfällen weit mehr, sind keine Ausnahme.[62] Einerseits nimmt die nötige Konzentration bei einer solchen Aufgabe schon nach weniger als einer Stunde rapide ab, andererseits können nicht alle verfügbaren Bilder mit gleicher Intensität ausgewertet werden.[63] OperateurInnen müssen daher Kriterien zur Filterung anwenden. Diese zweite Informationsselektion beruht auf expliziten oder impliziten Annahme der OperateurInnen über Risikoorte, Risikopersonen und Risikoobjekte bei der Überwachungstätigkeit.[64]
In Kapitel 5.2 wird untersucht, wie sich die Aufgabe der OperateurInnen durch Automatisierung ändert und welchen Einfluss dies auf die OperateurInnen hat.

2.3.2 Bedeutung von Datenschutz und *privacy*

Im Folgenden werden Datenschutz und *privacy* im Zusammenhang mit *manueller Videoüberwachung* betrachtet. Beide Begriffe werden in der Li-

61 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 124.
62 **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 3.
63 Ebd., S. 3.
64 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 125.

teratur nicht einheitlich definiert und interpretiert. Die Europäische Union versteht unter Datenschutz „insbesondere den Schutz der Privatsphäre natürlicher Personen bei der Verarbeitung personenbezogener Daten."[65] Das Sphärenmodell mit der Einteilung in Intimsphäre, Privatsphäre und Öffentlichkeitsphäre ist für das hiesige Begriffsverständnis jedoch obsolet geworden, als im Zusammenhang mit dem Volkszählungsurteil das Bundesverfassungsgericht feststellte, dass es keine „belanglosen Daten" gibt.[66] Auch ein für sich gesehen „belanglos" wirkendes Datum kann mit Verarbeitungs- und Verknüpfungsmöglichkeiten der Informationstechnik einen neuen Stellenwert bekommen.

Seit dem Volkszählungsurteil ist der Datenschutzbegriff außerdem aufs engste mit dem *Recht auf informationelle Selbstbestimmung* verbunden, das aus dem allgemeinen Persönlichkeitsrecht (Art. 2 Abs. 1, Art. 1 Abs. 1 GG) abgeleitet wurde. Es „gewährleistet das Recht des Einzelnen, grundsätzlich über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen".[67].

Im Bundesdatenschutzgesetz wird der Datenschutz daher als „Schutz des Persönlichkeitsrechts bei der Verarbeitung personenbezogener Daten" beschrieben.[68]

Auch der Begriff *privacy* wird auf unterschiedliche Weise genutzt und interpretiert.[69] Unter anderem ist – wie auch im Folgenden – das „Recht auf informationelle Selbstbestimmung" gemeint.

Alan Westin definierte *information privacy* als:

---

65 Art. 1 Abs. 1 Richtlinie 95/46/EG.

66 Bundesverfassungsgericht: Bundesverfassungsgerichtsurteil zur Volkszählung, 1 BvR 209, 269, 362, 420, 440, 484/83, BVerfGE 65, 1, 15.12.1983, S. 45.

67 **Dix**, „Datenschutz und Informationsfreiheit : Bericht 2010".

68 §1, Bundesdatenschutzgesetz (BDSG) vom 14. Januar 2003, zuletzt geändert am 14. August 2009 (BGBl. 2009 I S. 2814).

69 Charles Raab unterteilt den übergeordnete Begriff *privacy* beispielsweise in *privacy of the person*, *privacy of behaviour and action*, *privacy of personal communication*, *privacy of data and image*, *privacy of thoughts and feelings*, *privacy of location and space* und *privacy of association* (einschließlich *group privacy*).

> Das Recht von Individuen, Gruppen oder Institutionen selbst festzulegen, wann wie und in welchem Ausmaß, anderen Informationen über sich kommuniziert werden.[70]

Art und Umfang dieser Forderung variieren zwischen verschiedenen Gesellschaften, innerhalb einer Gesellschaft zeitlich, und zwischen einzelnen Individuen außerordentlich stark.[71] Für die Ausgestaltung von Videoüberwachungssystemen erweist sich dies als erschwerend.
Auch wenn die späteren Betrachtungen zu automatisierter Videoüberwachung sich nicht auf den deutschen oder europäischen Raum beschränken, sollen die Begriffe Datenschutz und „Recht auf informationelle Selbstbestimmung" (verkürzend als *privacy* bezeichnet) nach dem Verständnis des Bundesdatenschutzgesetzes (BDSG) verwendet werden. So bezieht sich auch die folgende Betrachtung zur Relevanz von Datenschutz bei Videoüberwachung auf das Bundesdatenschutzgesetz.

Relevanz von Datenschutz bei Videoüberwachung

Im Rahmen von Videoüberwachung erhobenes Bildmaterial und den darin enthaltenen Informationen sind datenschutzrelevant, da es sich im Sinne des Bundesdatenschutzgesetzes in der Regel um „personenbezogene Daten" handelt. Dabei ist unerheblich, ob tatsächlich eine Identifizierung vorgenommen wird oder vorgesehen ist, denn nach §3 Absatz 1 BDSG ist bereits die *Bestimmbarkeit* der Betroffenen ein hinreichendes Kriterium für „Personenbezogenheit" der Daten. Um dies zu betonen wird der Begriff „personenbeziehbar" statt „personenbezogen" verwendet. Aufgrund der interpersonell unterschiedlichen Ausprägung von *privacy* und der Tatsache, dass Daten nicht nach Relevanz unterschieden werden, sind potentiell alle im Rahmen einer Videoüberwachungsmaß-

70 **Westin**, *Privacy and Freedom*, (eigene Übersetzung).
71 **Tichy** ; **Peissl**, *Beeinträchtigung der Privatsphäre in der Informationsgesellschaft*, S. 8.

nahme erhobenen personenbeziehbaren Informationen relevant für den Datenschutz.

Informationsasymmetrie

Wie in Kapitel 2.3.3 beschrieben wird, herrscht zwischen Überwachern und Betroffenen auf Grund der *Distanz* ein starkes informationelles Ungleichgewicht (Informationsasymmetrie). Zusätzlich ist Betroffenen in den meisten Fällen unbekannt, welche Daten erhoben werden, aus welchem Grund sie erhoben werden, und wie sie *verwendet* d.h. verarbeitet und genutzt werden.[72] Modalitäten wie Speicherungsdauer, vorgesehene Zugriffsberechtigungen sowie Absicherung gegen Missbrauch und unberechtigten Zugriff[73] sind oft nicht bekannt bzw. können nicht ohne weiteres von Betroffenen überprüft werden.[74] Durch diese starke Informationsasymmetrie wird die Möglichkeit, das Recht auf informationelle Selbstbestimmung wahrzunehmen, stark eingeschränkt.

Der Verlust von *privacy* ist jedoch nicht das einzige Problem, das Videoüberwachung mit sich bringt. Im Folgenden werden eine Reihe weiterer Probleme und Auswirkungen beschrieben.

2.3.3 Wirkung auf Individuen und Gesellschaft

Extraktion und Projektion von Informationen

Kameras können nicht nur Informationen aus dem Raum extrahieren, sondern auch Informationen in den Raum projizieren.[75] Extraktion und Projektion können gut anhand zweier Extremfälle beschrieben werden:

72 Vgl. § 3 (5), Bundesdatenschutzgesetz (BDSG) vom 14. Januar 2003, zuletzt geändert am 14. August 2009 (BGBl. 2009 I S. 2814).

73 Bei analoger Übertragung kann die Datenübertragung leicht abgefangen werden. (Vgl. **Senior**, *Protecting Privacy in Video Surveillance*, S. 92).

74 Vgl. **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 352.

75 Ebd., S. 120.

versteckte Kameras einerseits und *Kameraattrappen* andererseits. Haben Betroffene keine Kenntnis über die Existenz einer Kamera (z. B. weil sie nicht sichtbar ist), wird von der Kamera keine Information in den Raum projiziert, sondern nur Informationen extrahiert, die ausgewertet und aufbewahrt werden können. Kameraattrappen hingegen können keine Informationen extrahieren, sondern können – sichtbar angebracht – Informationen in den Raum projizieren, somit in den Raum hineinwirken und das Handeln betroffener Personen in diesem Raum beeinflussen. Dieser Effekt kann auch unabhängig von der tatsächlichen Existenz von Kameras entstehen, wenn Betroffene glauben bzw. nicht ausschließen können, überwacht zu werden.

Distanz

Im öffentlichen Raum wird Videoüberwachung als Mittel zur sozialen Kontrolle eingesetzt. Im Allgemeinen findet soziale Kontrolle als ein Resultat der gleichzeitigen Anwesenheit und Überwachung der Menschen statt, die sich am gleichen Ort befinden.[76] Der Blick unbestimmter anderer stellt einen sozialen Regulationsmechanismus dar, der die Verhaltens- und Benutzungsnormen verschiedener Orte festigt.[77] Videoüberwachung findet jedoch aus der Ferne statt. Durch diese räumliche Entfernung und die in 2.3.1 beschriebene informationelle Kluft, wird die *Distanz* von Überwachung durch anwesendes Sicherheitspersonal zwischen Betroffenen und Überwachenden erhöht. Betroffenen wird es nicht ermöglicht, zu reagieren oder durch Entschuldigungen oder Erklärungen korrektive Informationen zu geben.[78] Durch diese fehlende Kommunikationsmöglichkeit erhalten gewisse Handlungen und Geschehnisse aus Sicht der

76 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 425.

77 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 170.

78 Ebd., S. 172.

OperateurInnen, die eine Risikosituation einschätzen sollen, eine „verstärkte »indikatorische« Bedeutung".[79] Darüber hinaus sind die genauen Umstände über Auswertung und Speicherung der Bilder oft unbekannt oder können von Betroffenen nicht überprüft werden. Zwischen Beobachteten und Beobachtenden herrscht daher, wie bereits erwähnt, eine starke Informationsasymmetrie.

Selbstdisziplinierende Wirkung und Konformität

Foucault entwickelte mit Bezug auf das von Bentham ausgearbeitete panoptische Prinzip eine Disziplinierungsthese.[80] Durch die umfassende Sichtbarkeit von Gefängnisinsassen bei gleichzeitiger Unsichtbarkeit des Wärters wissen die Gefangenen nie, wann sie beobachtet werden. Um Sanktionen zu entgehen, müssen sie die Kontrolle und damit die Herrschaft verinnerlichen. Unter der panoptischen Kontrolle disziplinieren sich die Gefangenen somit selbst.

Diese Effekte entstehen durch die beschriebene Informationsasymmetrie ebenfalls bei Videoüberwachung und wirken sich auf Individuen und langfristig auch auf die Gesellschaft aus. Wenn Personen eine Videoüberwachung vermuten oder sich der Überwachung bewusst sind, ist es wahrscheinlich, dass sie sich konform zur vermeintlichen Norm verhalten.[81] Eine konkretere Ausprägung dieser Disziplinierung beschreibt der *chilling effect* (engl. für abschreckende Wirkung), der aus der Vermutung erwächst, dass ein Staat, obwohl kein Fehlverhalten gezeigt wird, sich dazu entschließt, Daten aufzubewahren.[82] Diese Daten könnten später vom Staat oder einer späteren Form oder Instanz des Staates mit anderen Werten gegen die Bürger ausgenutzt werden. Zu wissen oder nicht aus-

79 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 172.

80 **Apelt** ; **Möllers**, „Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung", S. 589.

81 **Macnish**, „Surveillance Ethics", Abschnitt 4.

82 Ebd., Abschnitt 12.

schließen zu können, dass derartige Datensammlungen erhoben werden, könnte Menschen dazu disziplinieren, an legitimen Aktivitäten nicht teilzunehmen. Dieser *chilling effect* steht mit Menschenrechten im Konflikt, kann zu Konformität des Verhaltens und zum „Ersticken" von Kreativität führen. Vielfalt – als Gegenteil von Konformität – wird jedoch als eine Grundbedingung für Demokratie, Zivilgesellschaft und die gesellschaftliche-kulturelle und wirtschaftliche Entwicklung angesehen.[83]

Empirische Nachweisbarkeit

Theoretische Arbeiten deuten auf die Gefahr verstärkter Konformität hin. Der kontrollierte empirische Nachweis für den öffentlichen Raum ist jedoch problematisch, da es schwierig ist, den Anteil der Beobachtung durch Überwachungssysteme herauszuarbeiten.[84] Im vergangenen Jahr beschrieb jedoch Kirstie Ball Studien, die gesellschaftliche Kosten durch überwachungsbezogene Konformität in bestimmten sozialen Bereichen nachweisen konnten.[85] Überwachung am Arbeitsplatz bewirkte eine Reduzierung kreativen Verhaltens („reduction in creative employee behaviour") sowie eine Erhöhung von Stress, Widerstand und Sabotage. In Schulen störte sie die Qualität der Interaktion zwischen LehrerInnen und SchülerInnen bezüglich Harmonie, Vertrauen, Spontaneität und Kreativität. Außerdem war ein Verfall der Qualität der sozialen Interaktion zwischen SchülerInnen bezüglich der Form der Selbstdarstellung und Kreativität und in *gated housing projects* ein erhöhtes Bewusstsein der eigenen öffentlichen Erscheinung und des Verhaltens zu beobachten. Es wird außerdem festgestellt, dass Kommunikation innerhalb von Gruppen reduziert wird, Kontakt mit vermeintlich überwachten Gruppen oder Ide-

83 **Tichy** ; **Peissl**, *Beeinträchtigung der Privatsphäre in der Informationsgesellschaft*, S. 9.
84 **Apelt** ; **Möllers**, „Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung", S. 589,590.
85 **Ball**, „surveillancy and conformity", S. 233.

en vermieden und Überwachungsmaßnahmen mit Schuld und sozialer Unerwünschtheit der Überwachten assoziiert werden.[86]
Ein weiteres Problem bei *manueller Videoüberwachung* ist die Diskriminierung der Überwachten durch OperateurInnen.

Diskriminierung

In Kapitel 2.3.1 wurde beschrieben, dass OperateurInnen in der Regel mehr Videomaterial angezeigt wird, als sie tatsächlich auswerten können. OperateurInnen müssen daher nach irgendwelchen Kriterien auswählen, was einer näheren Beobachtung wert erscheint.[87] Untersuchungen in Kontrollräumen zeigen, dass die Sinnesbeschränkungen und die Distanz dazu ermutigen, Verdacht anhand einer begrenzten Palette einfach zu beobachtender Merkmale zu schöpfen, anstatt bei auffälligem Verhalten.[88] Eine Studie von Norris und Armstrong aus dem Jahr 1999 zeigte, dass anhand von Stereotypen, Alter und Geschlecht und besonders im Bezug auf Kriminalität und Terrorismus auch anhand von Ethnie und religiöser Identität ausgewählt wurde. 65% der Teenager standen ohne ersichtlichen Grund unter besonderer Beobachtung während das nur für 21% der über 30-Jährigen zutraf. Schwarze wurden doppelt so wahrscheinlich (68%) wie Weiße (35%) ohne ersichtlichen Grund beobachtet und Männer drei mal so häufig (47%) wie Frauen (16%).[89] Auch Kleidungsstil stellte sich als Kriterium heraus. Da diese Differenzierung nicht objektiv anhand von Verhalten und individuellen Charakteristika vorgenommen wurde, sondern mehrheitlich durch Kategorisierung zu einer bestimmten sozialen Gruppe, kann diese Praxis eindeutig als diskriminierend bezeichnet werden.[90]

86 **Raab**, „Impact of Surveillance on civil liberties and fundamental rights", S. 267.
87 **Norris** ; **Armstrong**, „CCTV and the social Structuring of Surveillance", S. 160.
88 **Hempel** ; **Töpfer**, „Videoüberwachung in Europa : Abschlussbericht", S. 8.
89 **Norris** ; **Armstrong**, „CCTV and the social Structuring of Surveillance", S. 163.
90 Ebd., S. 175.

Bei längerer Durchführung kann es zu Rückkoppelungseffekten kommen. Werden beispielsweise schwarze Teenager genauer beobachtet, so werden auch proportional mehr z. B. beim Diebstahl beobachtet und verurteilt. Die OperateurInnen werden in ihrer Annahme bestätigt und Kriminalitätsstatistiken werden verfälscht. Erfahrung und Statistiken scheinen dann wiederum den zukünftigen Fokus auf genau jene Gruppen zu rechtfertigen.[91] Diskriminierung wird auf diese Weise nicht nur reproduziert, sondern trägt zu ihrer Verfestigung und Verschärfung bei.[92]

Voyeurismus

Die Studie von Norris und Armstrong ergab außerdem, dass 15% der von OperateurInnen durchgeführten genaueren Beobachtung von Frauen aus voyeuristischen Motiven erfolgten.[93] Die Studie wird teilweise als veraltet und nicht repräsentativ angesehen. Folgestudien, welche die Ergebnisse relativieren, liegen jedoch noch nicht vor.[94]

Kriminalisierung

Neben der Diskriminierung besteht beim Einsatz von Videoüberwachung außerdem die Gefahr der Kriminalisierung. Welches Verhalten als Normalität wahrgenommen wird, unterscheidet sich von Gruppe zu Gruppe. Wenn eine bestimmte Verhaltenscharakteristik zur Identifikation einer Bedrohung definiert wird, entsteht das Risiko, dass eine Gruppe, die dieses Verhalten in einer nicht bedrohlichen Art zeigt, als Gefahr angesehen wird.[95] Bestimmte Kleidung und Symbole, können beispielsweise als Indiz für Gewaltbereitschaft oder Kriminalität betrachtet werden; eine körperliche Auseinandersetzung beispielsweise wandelt sich erst dann

91 **Macnish**, „Unblinking eyes : the ethics of automating surveillance“, S. 8.
92 **Apelt** ; **Möllers**, „Wie „intelligente“ Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung“, S. 590.
93 **Norris** ; **Armstrong**, „CCTV and the social Structuring of Surveillance“, S. 174.
94 **Macnish**, „Unblinking eyes : the ethics of automating surveillance“, S. 23.
95 Ebd., S. 9.

vom Konflikt zur Kriminalität, wenn jemand anderes es als „kriminell" benennt. Wird derartiges Auftreten mit Videoüberwachung verhindert oder an andere Orte verschoben, findet eine Stigmatisierung statt, die zu weiterer Kriminalität führen kann.[96] Der Effekt wird durch die Distanz und die fehlende Möglichkeit der Kommunikation verstärkt.

Reduziertheit des Diskurses auf Verletzung von Datenschutzbestimmungen

Das Spektrum der Auswirkungen von Videoüberwachung, das hier angeführt wurde, ist im öffentlichen Diskurs bisher noch nicht wiederzufinden. Die öffentliche Wahrnehmung der Problematik ist sehr einseitig. Die jüngst erschienene Diskursanalyse „Privacy issues in public discourse: The case of 'smart' CCTV in Germany" von Norma Möllers und Jens Hälterlein ergab zwar, dass Videoüberwachung hauptsächlich auf Grund der Spannung zwischen öffentlicher Sicherheit und persönlicher Freiheit als unangemessene Technik bewertet wurde. Sie zeigte jedoch, dass die Art und Weise wie der Begriff der *persönlichen Freiheit* im Diskurs definiert wurde, auf die Verletzung von Datenschutzbestimmungen reduziert war. Darin besteht die Gefahr, dass Videoüberwachung im Diskurs als legitim angesehen wird, sobald die Systeme Datenschutzbestimmungen einhalten würden.[97]

Weiterführende gesellschaftliche Auswirkungen

Videoüberwachung hat viele weiterführende, mitunter tief in das gesellschaftliche Leben eingreifende Auswirkungen, die weder vollständig vorhergesehen, noch empirisch nachgewiesen werden können. Als Beispiel sei hier der Effekt der sozialen Ausgrenzung[98] genannt – wenn Video-

96 **Albrecht**, *Der Weg in die Sicherheitsgesellschaft : Auf der Suche nach staatskritischen Absolutheitsregeln*, S. 20,21.

97 **Möllers** ; **Hälterlein**, „Privacy issues in public discourse: the case of "smart" CCTV in Germany", S. 10.

98 **Hempel** ; **Töpfer**, „Videoüberwachung in Europa : Abschlussbericht", S. 8.

überwachung sich z. B. gegen den Aufenthalt von Jugendgruppen oder Obdachlosen in Kaufhäusern richtet[99]. Ein anderes Beispiel sind die Bedenken, dass durch die Existenz der Videoüberwachung die Bereitschaft sozialer Akteure vermindert würde, bei Problemsituationen in öffentlichen Räumen selbst einzugreifen.[100] Auf diese weiteren Effekte kann hier nicht weiter eingegangen werden. Die Beispiele verdeutlichen jedoch die Notwendigkeit, über das Wesen und mögliche Auswirkung der Videoüberwachung nachzudenken und zu untersuchen, wie sich automatisierte Videoüberwachung auf solche Effekte auswirkt.

2.4 Zusammenfassung

In diesem Kapitel wurde das Problemfeld *Videoüberwachung* vorgestellt. Die Schilderung, dass trotz ausgebliebener Evaluation der Tauglichkeit weltweit ein massiver Ausbau von Videoüberwachung vorgenommen wurde, ohne dass dabei gesellschaftliche Auswirkungen und Grundrechte hinreichend berücksichtigt wurden, konnte in diesem Kapitel die Brisanz des Themenkomplexes aufzeigen. Es wurde beschrieben, wie diese Defizite sich weder in der Akzeptanz in der Bevölkerung noch im öffentlichen Diskurs wiederfinden und der Ausbau und die Forschung an Techniken trotzdem weiter vorangetrieben werden.

Zur Veranschaulichung der technischen Entwicklung wurden Videoüberwachungssysteme nach Grad der Automatisierung in drei Generationen eingeteilt und so die Begriffe *manuelle* und *automatisierte* Videoüberwachung unterschieden.

Ausgehend davon wurden Konzepte und Problembereiche *manueller Videoüberwachung* vorgestellt, die für die nachfolgende Identifizierung ge-

99 **Apelt** ; **Möllers**, „Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung", S. 589.

100 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 79.

sellschaftlicher Probleme automatisierter Videoüberwachung im öffentlichen Raum von Bedeutung sind. Der Begriff „öffentlicher Raum" wurde definiert. Er zeichnet sich nicht durch den juristischen Status des Raumes aus, sondern hauptsächlich durch die *soziale Begegnung*, die in ihm stattfindet.

Zur Herleitung der Probleme wurde zunächst die Struktur der Videoüberwachung und die Aufgabe der OperateurInnen beschrieben.

Videoüberwachung kann öffentliche Räume in ihrer Komplexität nicht erfassen und den OperateurInnen nicht immer alle Informationen zur adäquaten Einschätzung der Situation zur Verfügung stellen. Um diesem Informationsmangel entgegenzuwirken, müssen OperateurInnen die Informationen rekontextualisieren. Dabei spielen Erfahrungen der OperateurInnen und ihr Wissen über den überwachten Raum eine entscheidende Rolle.

Des Weiteren wurde dargelegt, dass Datenschutz bei Videoüberwachung von besonderer Bedeutung ist. Erhobene Daten sind durch Personenbeziehbarkeit datenschutzrechtlich relevant. Durch eine starke Informationsasymmetrie zwischen Überwachern und Betroffenen ist das Recht auf informationelle Selbstbestimmung stark eingeschränkt.

Auswirkungen, die über den Verlust von *privacy* hinausgehen, werden nicht nur in der Theorie vermutet, sondern konnten empirisch nachgewiesen werden. Neben Selbstdisziplinierung und Konformität kann Videoüberwachung auch zu Diskriminierung und Kriminalisierung führen. Nicht nur konkrete Maßnahme können individuelles Verhalten verändern. Videoüberwachung kann auch außerhalb des direkten Wirkungsbereichs einer Maßnahme Auswirkungen haben, die mitunter aber schwer empirisch nachweisbar sind. Videoüberwachung kann als Gefährdung von Zivilgesellschaft, Demokratie und Innovation und somit als Gefährdung gesellschaftlich-kultureller und auch wirtschaftlicher Entwicklung ange-

sehen werden. Der überwachungskritische Diskurs reduziert sich jedoch fast ausschließlich auf Verletzung von Datenschutzbestimmungen und lässt die anderen genannten Probleme bisher unberücksichtigt.

3 Techniken und Konzepte automatisierter Videoüberwachung

In short, the goal of visual surveillance is not only to put cameras in the place of human eyes, but also to accomplish the entire surveillance task as automatically as possible.[101]

WEIMING HU

In 2.2 wurde Videoüberwachung am Grad der *Digitalisierung* und der *Automatisierung* in drei Klassen unterteilt. Während manuelle Videoüberwachungssysteme der Generation II als „bessere Bewegungsmelder" maximal assistierend die Aufmerksamkeit der OperateurInnen auf simple Geschehnisse lenken konnte (z. B. das Übertreten einer vorher bestimmten Grenze wie in Abb. 3), soll automatisierte Videoüberwachung einen größeren Anteil der in 2.3.1 beschriebenen Aufgaben der OperateurInnen übernehmen oder den Menschen in seiner bewertenden und schlussfolgernden Funktion weitestgehend ersetzen.[102] Zusätzlich sollen Aufgaben, die ein Mensch nicht oder zumindest nicht effizient erfüllen kann, automatisiert ausgeführt werden.

Begriffe

In der Fachliteratur und in den Medien wird eine Vielzahl von Begriffen verwendet, hinter denen eine ganze Reihe von Konzepten, Ideen, Techniken und kompletten Systemen stehen, die alle gemein haben, die Überwachungsaufgabe ganz oder in Teilen zu automatisieren. Einige der Begriffe sind:

- *automated surveillance*[103]
- *algorithmic surveillance*[104]

101 **Hu** et al., „A survey on visual surveillance of object motion and behaviors".
102 Ebd., S. 334.
103 **Macnish**, „Unblinking eyes : the ethics of automating surveillance".
104 **Möllers** ; **Hälterlein**, „Privacy issues in public discourse: the case of "smart" CCTV in Germany".

- *intelligent in-situ security systems for public spaces*[105]
- *intelligent distributed surveillance system*[106]
- *second generation CCTV*[107]
- *smart CCTV*[108]
- *intelligente Videoüberwachung*[109]
- *intelligent video surveillance networks*[110]
- *intelligente Kamerasysteme*[111]

Möllers und Hälterlein diskutieren einige der Begriffe – keiner von ihnen wird als zufriedenstellend angesehen. *Algorithmic surveillance* unterscheide nicht zwischen visueller und nicht-visueller Überwachung und sei daher eher als Oberbegriff einzuordnen. Der Begriff *Second generation CCTV* reduziere komplexe, miteinander verwobene gesellschaftliche und technische Entwicklungen zu einer linearen evolutionären Logik technischer Entwicklung.
Die häufigste Verwendung in wissenschaftlichen Veröffentlichungen und Medien findet der Begriff *smart CCTV*. Der Begriff*Smart* – deutsch im Sinne von geschickt, elegant oder klug – muss jedoch wegen der starken normativen Konnotation eines technischen Fortschritts kritisiert werden. Auch die Bezeichnung *CCTV* (*closed circuit television*), ist obsolet,

105 **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces".
106 **Valera** ; **Velastin**, „Intelligent distributed surveillance systems: A review".
107 **Möllers** ; **Hälterlein**, „Privacy issues in public discourse: the case of "smart" CCTV in Germany".
108 Ebd.
109 Ebd.
110 **Coudert** ; **Dumortier**, „Intelligent Video Surveillance Networks: Data Protection Challenges".
111 **Winkler**, „Vertrauenswürdige Videoüberwachung : Sichere intelligente Kameras mit Trusted Computing".

da Systeme durch Vernetzung nicht mehr als „geschlossenes" (*closed*) System angesehen werden können.

Mehr als nur Auswertung von Videomaterial

Möllers und Hälterlein untersuchten das Spektrum der Kritik an „smart CCTV" im öffentlichen Diskurs. Dabei wurden nur Überwachungssysteme berücksichtigt, welche Videomaterial mit Techniken zur Mustererkennung analysieren und interpretieren. Wie im Folgenden gezeigt wird, wird jedoch an weit aus mehr als nur der Mustererkennung in Bewegtbildmaterial geforscht, um Videoüberwachung bzw. Raumüberwachung zu automatisieren.

Der Trend entwickelt sich dahingehend, Systeme möglichst flexibel zu entwerfen und neben Kameras auch beliebige andere Sensoren in die Systeme zu integrieren und die Systeme untereinander zu verknüpfen.[112] Außerdem sollen in Zukunft Informationen aus dem Internet und verschiedenen externen Datenbanken in die automatisierte Analyse und Bewertung von Geschehnissen und Situationen einfließen. Über die Speicherung der erhobenen Daten ist außerdem eine zeitlich rückwärts gerichtete Überwachung möglich. Die Überwachung geht in solchen Systemen also weit über die Echtzeitanalyse von Kamerabildern hinaus.

Um gesellschaftliche Probleme dieser neuen Systeme zu identifizieren, sind die ergänzenden Techniken zwingend in die Betrachtungen mit einzubeziehen. Im Folgenden wird daher nicht nur auf Bildverstehen eingegangen, sondern auch auf weiterführende Verarbeitung der aus den Bildern und anderen Datenquellen gewonnenen Informationen. Auch die Vernetzung, Kameratechnik sowie die Darstellung und Verwendung der Analyseergebnisse wird beschrieben. Die erwähnten anderen Datenquellen können hier nicht ausführlich betrachtet werden. Es sei jedoch mit Nachdruck darauf hingewiesen, dass durch die Verknüpfung der Video-

112 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy", S. 324.

überwachung mit Datenbanken und automatischer Analyse sozialer Netzwerke, Foren, Blogs, Fotodatenbanken oder Ähnlichem eine ganze Reihe neuer Auswirkungen und Probleme diskutiert werden müssen.

Struktur des folgenden Abschnittes

Im Folgenden wird ein Überblick gegeben, woran im Rahmen der Automatisierung von Videoüberwachung geforscht wird und welche Ansätze und Techniken bereits existieren. Die Reihenfolge in der diese vorgestellt werden, orientiert sich wowohl am Informationsfluss, der während des Einsatzes des Systems von Erhebung über Verarbeitung bis Abspeicherung oder Darstellung der Ergebnisse zu erwarten ist, als auch an der zu erwartenden Architektur eines möglichen Gesamtsystems, wie es exemplarisch in Kapitel 4 entworfen wird.

Dementsprechend wird in 3.1 zunächst die prinzipielle Funktionsweise der Analyse von beobachtbaren Verhalten, die zentraler Teil der Überwachungsaufgabe ist, beschrieben. Dazu werden die Abstraktionslevel beginnend mit basaler Bildverarbeitung auf Pixelebene bis hin zu höheren Abstraktionslevels des Bildverstehens und künstlicher Intelligenz vorgestellt. In diesen höheren Ebenen der Abstraktion werden Informationen gewonnen und für die Interpretation der Geschehnisse verwendet, die über visuell sichtbares *Verhalten* hinaus gehen. Diese weiteren Möglichkeiten Informationen aus den Bildern zu extrahieren werden in 3.2, 3.4.1 und 3.4.2 aufgeführt. Kameras stehen eigentlich am Anfang des Informationsflusses. Da sie jedoch vermehrt mit Rechenkapazität versehen werden und Teile der Bildverarbeitung auf ihnen geschieht, werden sie erst nach der Bildverarbeitung beschrieben. Zur Orientierung am Informationsfluss kommt dann der Aspekt der Struktur des Gesamtsystems hinzu. Nach den Kameras, die direkt an ein frei skalierbares Netzwerk angeschlossen werden können, werden weitere Netzwerkkomponenten wie Sensoren und Datenspeicher beschrieben. Der Logik des Informations-

flusses folgend werden dann Datenbanken (Kapitel 3.4.1) beschrieben, in denen gewonnene Informationen und Analyseergebnisse für eine spätere Auswertung abgelegt werden können. Zusätzlich zu den vom System selbst erhobenen Daten, können auch externe Datenquellen, auf die über das Internet zugegriffen werden kann oder andere Datensammlungen bei der Analyse verwendet werden (3.4.2).
An letzter Stelle der Prozesskette stehen die OperateurInnen, die in ihrer Funktion als Teil des Systems betrachtet werden müssen. Daher werden anschließend in 3.5 Ansätze der Visualisierung der Analyseergebnisse für die Mensch-System-Interaktion beschrieben.
Ein gesonderter Abschnitt (3.6) wurde Techniken gewidmet, die der Bewerkstelligung von *privacy* dienen sollen.
Da der Charakter automatisierter Videoüberwachung nicht nur durch die Funktionsweise der einzelnen Techniken geprägt ist, sondern erst anhand ihres Zusammenspiels deutlich wird, wird in Kapitel 4 ein wahrscheinliches Komplettsystems entworfen, an dem in 5 gesellschaftliche Probleme identifiziert werden sollen.

3.1 Verhaltenserkennung

Ein Ziel von Videoüberwachung im *öffentlichen Raum* ist die Aufrechterhaltung oder Herstellung eines gewünschten Zustandes des überwachten Raumes. Mit der Echtzeitauswertung der Bilder wird bezweckt, Umstände zu registrieren, die von diesem Zustand abweichen oder vermutlich in Zukunft zu einer Abweichung führen.
Da meist menschliches Handeln diesen Zustand gefährdet oder aber die Unversehrtheit von Menschen selbst Teil des gewünschten Zustandes ist, steht die Auswertung menschlichen Verhaltens im Mittelpunkt. Welche Konzepte und Techniken für diese Aufgabe entworfen wurden, wird im Folgenden beschrieben.

3.1.1 Verhalten und Erkennungsansätze

In der *Computer Vision* wird der Begriff *Verhalten* in seinem allgemeinsten Sinn verstanden, nämlich als die beobachtbaren Aktionen von „Agenten". Diese Agenten können Menschen, Tiere, Autos oder andere Objekte sein.[113]

Blacklist- und *Whitelistansatz*

Zur Detektion von unerwünschtem oder auffälligem Verhalten lassen sich zwei grundsätzliche Varianten unterscheiden,[114] die als *Blacklist-* und *Whitelistansatz* bezeichnet werden sollen. Der *Whitelistansatz* entspricht der Modellierung des gewünschten Zustandes und dem Messen von *Abweichungen* von diesem Zustand. Es wird dafür eine Menge von geduldeten oder erwünschtem Verhalten modelliert. Während der Überwachung wird von diesen Modellen abweichendes oder Verhalten, dass zu keinem der Modelle passt, als deviant eingestuft. Wird als normal z. B. nur „übliches Gehen" modelliert, bewertet das System Hüpfen, Kriechen oder Ducken als abnormal.[115]

Der *Blacklistansatz* entspricht der expliziten Modellierung unerwünschten Verhaltens und der Messung von *Übereinstimmung* mit dem Modell. Eine Schlägerei könnte beispielsweise modelliert werden als schnelle Bewegungen einer Hand gegen den Körper einer anderen Person. Fußtritte würden dann nicht als Schlägerei erkannt werden, eventuell aber das „Rückenklopfen" beim Helfen einer hustenden Person.

Modellbegriff und Form, in der die Modelle vorliegen

Die Begriffe „Modell" und „Modellierung" verleiten zur Annahme einer gewissen Abgeschlossenheit und Eindeutigkeit dieser Verhaltensrepräsen-

113 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?", S. 333.
114 **Wiliem** et al., „A Context-Based Approach for Detecting Suspicious Behaviours", S. 1.
115 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?", S. 335.

tationen. Um dieser falschen Vorstellung für die folgenden Erklärungen vorzubeugen, muss an dieser Stelle technisch ein wenig vorweggegriffen werden. Die Form, in der die Modelle nämlich vorliegen, kann sehr komplex sein. Einzelne Aspekte des Verhaltens, die nach dem menschlichen Verständnis semantisch untrennbar zusammengehören, müssen mit unterschiedlichsten Techniken erfasst, auf unterschiedlichsten Abstraktionsebenen des Systems verarbeitet, in unterschiedlichsten Datenstrukturen repräsentiert sein. Die Zahlenwerte die ein Verhalten repräsentieren, können daher untrennbar verwoben mit anderen Modellen und selbst für die EntwicklerInnen vollkommen unbestimmbar sein. Während etwa Trajektorien von typischen Laufwegen als Zahlenreihen vorliegen also nahezu menschenlesbar sind und verständlich visualisiert werden können, ist es mitunter kaum möglich, automatisiert gewonnene Modelle z. B. von Aggressivität zu beschreiben und noch viel weniger sie so darzustellen, dass sie eingehender untersucht werden können. Bei neuronalen Netzen, auf die weiter unten eingegangen wird, liegen die Modelle beispielsweise codiert in den gelernten Gewichten und der Topologie des Netzes vor. Man kann von einzelnen Neuronen oder Parametern komplexerer Strukturen kaum Aussagen über deren Rolle für das Gesamtnetzwerk bzw. über einzelne „Modelle" treffen. Im Folgenden sollen diese Eigenschaften von Verhaltensrepräsentationen bei der Benutzung des Begriff „Modell" mitgedacht werden.

Abstraktionslevel von Verhalten

Techniken, die „Auffälligkeit" auf einer niedrigen Abstraktionsebene detektieren, werden bereits eingesetzt. Kommerzielle Überwachungssysteme werden mit Algorithmen beworben, die beispielsweise ein Entfernen oder Hinzufügen eines Gegenstandes oder das Betreten eines bestimmten Bildbereichs erkennen. Einsatzgebiete sind z. B. Bahnhöfe (vgl. Abb. 3) oder die Überwachung von Ausstellungsgegenständen in Museen.

Von Verhalten kann hier nicht wirklich die Rede sein, vielmehr könnte man diese Funktionen als „bessere Bewegungsmelder" bezeichnen. Derartige Techniken werden in Systemen der Generation II eingesetzt, um den OperateurInnen durch Lenkung der Aufmerksamkeit zu assistieren. Techniken, die Daten auf einer höheren Abstraktionsebene verarbeiten, zum Beispiel „abnormale" Verhaltensmuster oder Verhalten verdächtiger Personen detektieren, sind immer noch experimentell.[116]

Zu repräsentierende Konzepte von Verhalten und Geschehnissen

Konzepte, die vom System repräsentiert und erkannt werden sollen, sind *basale Merkmale* wie Bewegungstrajektorien und Geschwindigkeit von Agenten und Objekten, deren *Zustände* zu einem bestimmten Zeitpunkt oder von bestimmter Dauer (z. B. eine Person ist „aufgeregt"), *Ereignisse*, die einen Zustandswechsel beschreiben und *Szenarien*, die eine Kombination aus Zuständen, Ereignissen und Subszenarien sind.[117] Explizit genannt werden hierzu Beispiele wie „Vandalismus", „Überfüllung von Orten", „Kämpfe", „unbeaufsichtigtes Gepäck" oder „Herumlungern". Es wurde außerdem erkannt, dass für die Bewertung von Verhalten und Geschehnissen auch ihr Kontext in Betracht gezogen werden muss.[118] Eine geringe Abweichung in dem einen Kontext kann auf Anomalie hinweisen, in einem anderen jedoch unverdächtig sein. Dazu wird auch ein Modell des Kontextes der Geschehnisse erstellt. Für die Kathegorisierung von Laufwegen könnte ein solches Modell beispielsweise übliche Ein- und Ausgänge eines Raumes beschreiben, also Bildbereiche, an denen Menschen erscheinen oder abtreten.[119] Von besonderem Interesse für Betreiber von Videoüberwachung sind Modelle, mit denen von beobachteten

116 **Raab**, „Impact of Surveillance on civil liberties and fundamental rights", S. 255.
117 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications", S. 287.
118 **Wiliem** et al., „A Context Space Model for Detecting Anomalous Behaviour in Video Surveillance".
119 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?", S. 332.

Verhalten auf zukünftige Geschehnisse und Verhalten geschlossen werden kann und so eine *vorwärts gerichtete Videoüberwachung* möglich erscheint.

3.1.2 Hierarchische Organisation

Computer-Vision-Systeme sind meist hierarchisch organisiert.[120] Bei der Auswertung von Videomaterial müssen letzlich aus einer Menge von Pixeln (unterste Ebene) Verhaltensinterpretationen auf höherer Abstraktionsebene berechnet werden. Meist findet dazu eine unidirektionale Verarbeitung der Daten statt. Typische Schritte sind Bewegungs- und Objekterkennung, Objektklassifizierung, Verdeckungsfolgerung und Objektverfolgung, Szenenmodellierung, Verhaltensanalyse und Erkennen eines bestimmten Ereignisses.[121] Die Bildverarbeitungstechniken auf niedriger Ebene speisen Trackingalgorithmen, die wiederum Szenen- und Verhaltensanalysemodule speisen. Ist ein Verhalten modelliert und liegt in numerischer Repräsentation vor, kann eine ganze Reihe statistischer Methoden dazu genutzt werden, das Verhaltensmodell zu klassifizieren.[122] Klassen von Verhalten sollen dann mit einer Semantik versehen und das Verhalten darauf basierend ausgewertet werden oder mit anderen Informationen in Relation gesetzt werden.
Die einzelnen Hierarchieebenen können grob eingeteilt[123] werden in

- Bewegungsdetektion

120 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 331.

121 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications“, S. 297.

122 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 331.

123 Bei*Bewegungsdetektion*, *Objektverfolgung* und *Verhaltensmodellierung* wird sich an einer Studie von Weiming Hu et al. orientiert. Um die Übersichtlichkeit zu wahren, wird diese Quelle im Folgenden nur abschnittsweise angegeben. Abweichende Quellen werden stets gekennzeichnet.

- Objektverfolgung und
- Verstehen von Verhalten.

1) Bewegungsdetektion

Ausgangsdaten sind zeitliche Folgen von zweidimensionalen Bildern, die aus Pixeln zusammengesetzt sind. Ziel der Bewegungsdetektion (*motion detection*) ist die Segmentierung von relevanten Regionen in Bildmaterial (*region of interest – ROI*), die zu einem sich bewegenden Objekt gehören. Dazu gehören die Modellierung der Umwelt, die Segmentierung der Bewegung und die Objektklassifizierung.

Modellierung der Umwelt

Bei der Modellierung der Umwelt wird auf Pixelebene der Hintergrund erkannt. Es gibt eine ganze Reihe von Algorithmen, die diese Aufgabe auch mit unter schwierigeren Umstände wie Hintergrundbewegungen (z. B. im Wind wehende Bäume), Helligkeitsschwankungen und Schattenwurf erfüllen. Dazu werden temporäre Durchschnitte von Bildsequenzen genutzt. Bei beweglichen Kameras werden temporäre Modelle für den Hintergrund genutzt. Neben dieser zweidimensionalen Herangehensweise existieren auch erste Versuche, dreidimensionale Umweltmodelle zu nutzen – wegen der Komplexität jedoch bisher nur in Innenbereichen. Als Annäherung an 3D-Modellierung können Tiefenkarten des Bildes erstellt werden, die für jeden Pixel den Abstand vom Hintergrund zur Kamera beschreiben.[124]

Segmentierung der Bewegung

Für die Segmentierung der Bewegungen werden für verschiedene Methoden räumliche und zeitliche Informationen genutzt. Bei statischen Szenen

124 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 332.

kann der Unterschied zwischen vorher modelliertem Hintergrund und aktuellem Bild berechnet werden (*background subtraction*). Dynamischer ist es jedoch pixelweise die Unterschiede zwischen zwei oder mehreren aufeinander folgenden Bildern zu bestimmen (*temporal differencing*). Die veränderten Bildbereiche, die evtl. noch Löcher enthalten, werden dann zu Bewegungsregionen zusammengefasst. Rechnerisch aufwendiger, aber vom Ergebnis her gewinnbringender, sind *optical-flow*-Methoden. Es wird berechnet, welche Matrix von Vektoren ein Bild in das darauffolgende, meist nur geringfügig veränderte Bild verzerren.[125] Diese Bewegungsvektoren werden dann zur Segmentierung genutzt. Hybridalgorithmen aus Varianten der drei genannten Methoden können den Rechenaufwand so reduzieren, dass sie auch Anforderungen der Analyse in Echtzeit genügen. In diesem Stadium können die ermittelten bewegten Regionen unter Umständen zu verschiedenen sich bewegenden Objekten gehören.

Objektklassifizierung

Die Klassifizierung dieser Objekte kann als Standardproblem der Mustererkennung betrachtet werden. Momentan ist der der Ansatz Objekte entweder basierend auf ihrer Form (*shape-based classification*) oder basierend auf ihrer Bewegung (*motion-based classification*) zu klassifizieren. Verschiedene Formen können mit Punkten, umrahmenden Boxen, Silhouetten oder sogennannten *blobs* (engl. für Klecks) beschrieben werden. Anhand von Verteilung, Seitenverhältnis der *bounding box* sollen die Formen in verschiedene Klassen von Objekten wie etwa Mensch, Fahrzeug, Menschengruppe oder sonstiges „Wirrwarr" (*clutter*) eingeteilt werden. Auch für die Objektklassifizierung wurden Hybridtechniken entwickelt.[126]

125 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications", S. 282.
126 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 337.

2) Objektverfolgung

Während bisher die Informationen über Bewegungen nur zur Detektierung und Klassifizierung von Objekten genutzt wurden, werden nun die Bewegungen dieser Objekte selbst von Bild zu Bild verfolgt (*object tracking*), um deren Ort im Raum und eventuell deren Pose zu bestimmen. Auch hierzu existiert eine ganze Reihe von Ansätzen, die sich kombinieren lassen. Man unterscheidet *active contour-based*, *feature-based* und *model-based tracking*.[127]

Active contour-based tracking

Beim *active contour-based tracking* werden die Umrisse des Objekts als einhüllende Kontur repräsentiert und in den aufeinander folgenden Bildern immer wieder aktualisiert. Diese Methode hat einen geringen Rechenaufwand und bleibt auch stabil bei unruhigen Szenen und bei partieller Verdeckung. Die Extraktion dreidimensionaler Posen eines Objekts im Raum sind mit diesem Verfahren eine Herausforderung.

Feature-based tracking

Feature-based tracking nutzt besondere visuelle Merkmale von Objekten, sogenannte *features*. Einzelne Objektelemente werden extrahiert und zu *features* zusammengefasst, die dann zwischen den aufeinander folgenden Bildern abgeglichen werden. Es werden globale *features* wie Schwerpunkt, Umfang, bestimmte Gebiete oder Farbe, sowie auch lokale *features* wie Liniensegmente, Wölbungen oder Eckpunkte genutzt. Auch geometrische Eigenschaften zwischen den *features* werden berücksichtigt. Abgesehen von Letzterem ist *feature-based tracking* ressourcensparend und wird daher auch zur Echtzeitverarbeitung eingesetzt. Die tatsächliche Erkennungsrate ist jedoch gering, da dieses Verfahren nur im Zweidimensio-

127 Ebd., S. 338.

nalen arbeitet und mit Variationen des Blickwinkels auf ein Objekt an Stabilität verliert.

Model-based tracking

Model-based tracking verfolgt Objekte mit Hilfe vorher konstruierter Modelle. Es wird die Herangehensweise *analysis-by-synthesis* genutzt nach dem Prinzip *predict-match-update*. Dabei wird die Pose des Modells für das nächste Bild anhand des bisherigen Verlaufs und anderen Wissens prognostiziert. Dann wird das Modell synthetisiert und auf die Bildebene projiziert. Die Projektion wird dann mit den tatsächlichen Bilddaten verglichen und die Pose des Modells gegebenenfalls angepasst und aktualisiert. Die Technik wird sowohl für *tracking* von Menschen als auch von starren Objekten genutzt. Für die Modellierung menschlicher Körper werden verschiedene Abstraktionen genutzt. Diese reichen von *stick figures* (Kopf, Rumpf und Gliedmaßen werden mit Stäben und Gelenken repräsentiert) über *volumetric models* (Zusammensetzung mehrerer einfacher geometrischer Körper) bis hin zu komplexeren Modellen wie *hierarchical models* bei denen Skelett, Fett und Oberfläche modelliert werden. Hier besteht ein Trade-off zwischen Präzision und Rechenaufwand.
Nach dem Bewegungen von Agenten und Objekten erfasst sind, soll nun eine semantische Auswertung dieser Daten erfolgen.

3) Modellierung und Verstehen von Verhalten

Die meisten der bis zu dieser Abstraktionsebene eingesetzten Techniken stammen aus den Forschungsgebieten *Bildverarbeitung* und *Computer Vision*. Die im Folgenden vorgestellten Verfahren sind der *Künstlichen Intelligenz* zuzuordnen.[128] Sie verarbeiten keine Bilder mehr, sondern Zahlenwerte, die – sich über die Zeit ändernde – Merkmale von Objekten repräsentieren. Diese sind die im vorherigen Schritt erhobene Raum-Zeit-

128 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications", S. 286.

Vektoren, Koordinaten oder Modellposen. In 3.2 wird besprochen, welche zusätzlichen Informationen aus dem Bildmaterial gewonnen und im nächsten Schritt berücksichtigt werden sollen. Die Aufgabe der folgenden Ebene besteht in der Klassifizierung der Informationen. Dazu müssen die erhobenen Bewegungsinformationen mit Verhaltens*modellen verglichen* werden. Auch hier müssen, auf Grund der Komplexität menschlichen Verhaltens, verschiedene Abstraktionsebenen gefunden werden, die von primitive Bewegungen bis hin zu komplexeren Szenen mit Interaktion zwischen den beobachteten Objekten und *Agenten* reichen.

Primitive Bewegung einzelner Objekte

Ein erster Schritt ist die Wiedererkennung bestimmter Bewegungsabläufe. Die meisten Ansätze verfolgen das Ziel, die Bewegungssequenzen in diskrete repräsentative Zustände zu segmentieren und diese Zustandssequenzen mit Referenzsequenzen unter Berücksichtigung einer gewissen Abweichung zu vergleichen. Es soll hier nur eine Auswahl gängiger Ansätze beschrieben werden:

Für die Analyse von Laufwegen kann eine *Trajektorienklassifizierung* durchgeführt werden.[129] Bewegungstrajektorien, die in Richtung und Verlauf nicht einer vorher festgelegten Klasse zuzuordnen sind, sollen erkannt werden.

Mit *Dynamic Time Warping* sollen Zustandssequenzen unabhängig von ihrer zeitlichen Entwicklung aufeinander abgebildet und ein Differenzmaß angewendet werden. Diese Technik wird vor allem zur Erkennung von Gesten eingesetzt. Auch *endliche Automaten* mit Zuständen, Zustandsübergängen und Zustandsübergangsfunktionen kommen für die Messung von Übereinstimmung mit Referenzsequenzen zum Einsatz. Am häufigsten genutzt werden *Hidden Markov Modelle* (HMMs). Ein *HMM* ist

129 **Khalid**, „Motion-based behaviour learning, profiling and classification in the presence of anomalies".

eine Art *stochastischer* Zustandsautomat mit bestimmten Wahrscheinlichkeiten der Zustandsübergänge. Auch sie sind flexibel bezüglich der variierenden Dauer von Bewegungsausführungen.[130] Bei der Modellierung von Bewegungsprimitiven mit *HMMs* müssen diese nicht explizit angegeben werden. Bewegungstrajektorien werden in Zustände zerlegt, die zum Beispiel durch Gelenkstellungen und Winkelgeschwindigkeiten charakterisiert sind. Die Übergangswahrscheinlichkeiten geben die Wahrscheinlichkeiten eines Zustandswechsels von einem Zeitpunkt zum nächsten an. Durch die Kombination von *HMMs* einzelner Bewegungsprimitiven sollen komplexe Bewegungen und Bewegungssequenzen zusammengesetzt werden. Die Folge von *HMMs*, welche die Bewegungstrajektorie am besten beschreibt, wird ermittelt und als die erkannte Bewegung ausgegeben.

Komplexere Geschehnisse

Komplexere Zusammenhänge wie Interaktion zweier Agenten oder eines Agenten mit einem Objekt (z. B. eine Person schiebt einen Einkaufswagen) sollen mit kombinierten *HMMs* (*coupled HMMs*) modelliert werden. Es gibt außerdem Ansätze zu Erkennung und Analyse zyklischer Bewegungen (*cyclic HMMs*) und Bewegungen mehrerer zusammengehörender Objekte wie z. B. zwei gestikulierende Hände (*parallel HMMs*).[131] Basierend auf *HMMs*, die primitive Ereignisse modellieren, sollen komplexere Ereignisse auch mit *syntaktischen* Techniken beschrieben werden.[132] Sie sollen mit Hilfe stochastischer oder auch nichtprobabilistischer *kontextfreier Grammatiken* beschrieben und von einem Parser erkannt werden. Hier findet ein Übergang von auf Wahrscheinlichkeiten basierender Modellierung zu Modellierung basierend auf binären Wahrheitswerten statt.

130 **Gehring**, **Kühne** ; **Schultz**, „Erkennung von menschlichen Bewegungen mit Hidden Markov Modellen“, S. 1 f.

131 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 333.

132 Ebd., S. 334.

Die Einführung solcher Abstraktionsebenen soll das Kombinieren basaler Bewegungen zu domainspezifischen Verhaltensmodellen erlauben. Es ist somit nicht mehr nötig, Verhaltensmodelle für jedes Einsatzgebiet (*domain*) von Überwachungsmaßnahmen von Grund auf neu zu erzeugen.

Neuronale Netze

Als letzter Ansatz sollen *Neuronale Netze* genannt werden. Im Einsatz werden dem Netzwerk an den Eingangsneuronen Zahlenwerte präsentiert. Jedes Neuron im Netz berechnet anhand der anliegenden Werte einen neuen Wert der über gewichtete Verbindungen an nachfolgende Neuronen propagiert wird. An den Ausgangsneuronen entstehen Werte, die abgegriffen werden können. Während einer Trainingsphase wird dem Netzwerk eine Reihe von Werten präsentiert. Abhängig von der jeweiligen Ausgabe können dann Neuronen hinzugefügt oder entfernt und Verbindungen, Schwellwerte und Wichtungen dahingehend angepasst werden, dass das Netz bei bestimmten Eingaben gewünschte Ausgabewerte erzeugt. Training und vor allem auch das Finden einer der Aufgabe angemessenen Ausgangstopologie des Netzes sind nicht trivial. Ein Beispiel für den Einsatz sind *Time-delay neuronal networks*. Diese wurden beispielsweise erfolgreich zum Lippenlesen oder zur Erkennung von Gesten angewendet.

Abweichungen von beschriebener Hierarchie

Die beschriebenen Hierarchieebenen dienen hier nur zur Veranschaulichung der grundsätzlichen Funktionsweise. Es gibt Ansätze, die nicht strikt unidirektional sind, sondern Informationen an untere Ebenen weitergeben oder Ebenen auslassen. Höhere Ebenen können so wiederum niedrige Prozesse der Informationsgewinnung beeinflussen und so je nach Bedarf an fehlende Informationen gelangen.[133]

133 **Yingjie** ; **Yin**, „Towards Suspicious Behavior Discovery in VideoSurveillance System", S. 541.

3.1.3 Gewinnung von Verhaltensmodellen

Referenzmodelle, mit denen das überwachte Verhalten verglichen wird, werden grundsätzlich auf zwei verschiedene Arten gewonnen. Sie werden *manuell* erstellt oder mit Verfahren maschinellen Lernens *automatisiert* erzeugt. Der Trend entwickelt sich hin zur automatisierten Erzeugung.[134]

Manuelle Erstellung

Manuelles Modellieren und Markieren von Verhaltensmustern ist sehr arbeitsaufwendig oder durch eine zu große Menge unterschiedlichen Verhaltens, das zu modellieren wäre, sogar unmöglich.[135] Außerdem wird mitunter angenommen, dass manuell erstellte Modelle inkonsistent und fehleranfällig sein können, da Menschen beim Modellieren dazu neigen, A-priori-Wissen mit einzubeziehen, das nicht visuell detektierbar ist. Um die Erstellung auch den Betreibern des Systems zu ermöglichen, ist die Modellierung mit Elementen und Konzepten der höheren Abstraktionsebenen zu erwarten. Denkbar ist die Beschreibung komplexerer Handlungen, die aus einzelnen Modulen eines Kataloges von Verhaltens- oder Bewegungskonzepten zusammengesetzt werden.

Automatisierte Erstellung

Bei automatisierter Modellierung werden dem System in einer Anlernphase Beispiele aus einem Trainingsdatensatz präsentiert in denen es automatisch Gesetzmäßigkeiten finden soll. Im Einsatz sollen diese Gesetzmäßigkeiten auch auf Datensätze angewendet werden, die von den Beispielen im Trainingsdatensatz abweichen. Man unterscheidet zwischen *beaufsichtigtem* und *unbeaufsichtigtem* Lernen.

134 **Valera** ; **Velastin**, „Intelligent distributed surveillance systems: A review", S. 196.
135 **Xudong**, **Hui** ; **Zhijing**, „Behavior Clustering for Anomaly Detection", S. 280.

Beaufsichtigtes Lernen

Beim beaufsichtigten Lernen müssen Paare von Ein- und dazu gegebenen Ausgaben zur Verfügung gestellt werden. Eingabewerte werden dem System präsentiert. Der Lernalgorithmus passt dann Beispiel für Beispiel ein mathematisches Modell so an, dass die mit dem Modell errechneten Werte möglichst mit den gewünschten Ausgabewerten des Trainingsdatensatzes übereinstimmen. Dafür werden Anhand der berechneten Unterschiede zwischen tatsächlichen und erwarteten Ausgabewerten automatische Modifikation von Parametern des Modells getätigt die den Unterschied minimieren. Diese Herangehensweise eignet sich aus technischer Sicht nur für Situationen, in denen von wohl-definierten und vorher bekannten Verhaltenskategorien ausgegangen werden kann.[136] Beim praktischen Einsatz kann es jedoch leicht zu einer so großen Menge unterschiedlichen Verhaltens kommen, dass beaufsichtigtes Lernen keine zielführende oder effiziente Lösung darstellt.

Unbeaufsichtigtes Lernen

Bei unbeaufsichtigtem Lernen wird versucht, Muster in Eingabedaten zu erkennen, ohne dass Vorgaben benötigt werden.

Unbeaufsichtigtes Lernen wird zur Komprimierung mittels Dimensionsreduktion und automatischer Segmentierung (*clustering*) eingesetzt. Das Ziel von *Komprimierung* ist, eine Menge von Daten in kompakterer Form darzustellen, in dem nur die Hauptkomponenten gespeichert werden und als unwichtig erachtete Daten weggelassen werden.

Beim automatischen *clustering* werden automatisiert Merkmale zur Unterscheidung gefunden, mit denen sich Muster in vorher nicht definierte Gruppen (*cluster*) einteilen lassen. Nach diesem Prinzip können auf nied-

136 Ebd., S. 281.

riger Abstraktionsebene z. B. Laufwege von Menschen automatisiert zu Gruppen üblicher Laufwege zusammengefasst werden.[137]
Auch auf höheren Abstraktionsebenen wird *clustering* eingesetzt. Ziel von *automatic behaviour clustering* ist, aus einer Sammlung nicht interpretierter Videos Modelle zu lernen, die in der Lage sind, vorher nicht aufgetretene, abnormale Verhaltensmuster, z. B. mutwilliges Zerkratzen von Autos auf einem Parkplatz zu detektieren und gleichzeitig neue Instanzen von „normalem", erwarteten Verhalten zu erkennen.[138] Somit sollen Modelle *online* – d.h. während des Einsatzes der Videoüberwachung – an Hand von neuem beobachteten Verhalten immer wieder angepasst und erweitert werden. Diese Herangehensweise entspricht dem *Whitelistansatz*, bei dem „Normalität" definiert wird und Abweichungen gemessen und als deviant eingestuft werden.

3.2 Weitere Möglichkeiten des Bildverstehens

Neben den Bestrebungen Verhalten und Geschehnisse zu erkennen, gibt es eine Reihe weiterer Bemühungen, auch andere Informationen aus dem überwachten Raum bzw. den Videomaterial automatisiert zu extrahieren, die für die Interpretation der Szene genutzt oder aus anderen Gründen gesammelt werden sollen. Im Folgenden soll ein kurzer Überblick über derartige Techniken und Ansätze gegeben werden.

3.2.1 Objekterkennung

Objekte sollen anhand eines Sets von visuellen Merkmalen (*features*) beschrieben werden und Anhand von genügend ähnlichen im Bild gefundenen *features* detektiert werden.[139] Auf diese Weise sollen auch Symbole z.

137 **Khalid**, „Motion-based behaviour learning, profiling and classification in the presence of anomalies".
138 **Xudong**, **Hui** ; **Zhijing**, „Behavior Clustering for Anomaly Detection", S. 297.
139 **Senior**, *Protecting Privacy in Video Surveillance*, S. 54.

B. die Beschriftung eines T-Shirts im Bild erkannt werden. Eine weitere Möglichkeit ist das Modellieren konkreter Objekte und der Abgleich des Beobachteten mit diesen Modellen. Dieser Ansatz ist vergleichbar mit dem bereits beschriebenen modellbasierten Ansatz beim Erkennen von Menschen. Bei Videoüberwachung kommt Objekterkennung bereits bei der Erkennung von unbewegten Gegenständen auf Fahrbahnen oder Wartekolonne vor Ampeln zum Einsatz.[140] In der Schweiz wurden 1998 in einem Pilotprojekt Nummernschilder erfasst und Alarme generiert, wenn in der Fahndung befindliche Autos erkannt wurden oder solche, für die nicht die obligatorische Haftpflichtversicherung bezahlt wurde.[141] Zielstellung des Forschungs-Projekt *INDECT* ist unter anderem, gefährliche Gegenstände wie Waffen in den Videobildern zu detektieren.[142]

3.2.2 Bewegung und Zusammengehörigkeit

Es gibt außerdem eine Reihe von Ansätzen, weitere Informationen aus den Bewegungen von Objekten und Personen zu gewinnen.
Um potentiell gefährliche Gegenstände zu entdecken, ohne dass sie aus Kameraperspektive sichtbar sind, sollen Ausweichbewegungen von Menschenströmen ausgewertet werden.[143] Es werden außerdem Techniken beschrieben, die Prognosen von Sichtbarkeit und von Verdeckungen ermöglichen sollen.[144] Dafür werden z. B. Sichtachsen zwischen Personen berechnet, an denen ermittelt werden soll, ob eine Person z. B. dem Sicherheitspersonal beständig aus dem Weg geht. Dies soll unter anderem auch mit drucksensitiven Bodenplatten realisiert werden. Zusätzlich sol-

140 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 60 f.
141 Ebd., S. 61.
142 **Rutz**, *Interview mit INDECT-Projekt-Koordinator*.
143 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 333.
144 Ebd., S. 333.

len auch Beschleunigung oder eine abrupte Änderung der Laufrichtung erkannt und ausgewertet werden. In Kombination mit Objekterkennung sollen mit *abandoned object's owner detection* allein gelassene Gegenstände wie etwa Gepäck erkannt werden und anhand der Videodaten der Besitzer oder die Besitzerin ermittelt werden. Aufgrund von Überdeckungen oder ungünstigen Kamerawinkeln kommt es dabei noch zu einer sehr hohen Rate von Falschalarmen und auch das Merken von BesitzerInnen ist noch problematisch.[145]

3.2.3 Mimik, Gestik und Körpersprache

Bereits erwähnt wurde das Erkennen von Gestik und Lippenlesen mit Hilfe neuronaler Netze.[146] Mit dem Ziel „Emotionen" zu erkennen gibt es eine ganze Reihe weiterer Forschungen.[147] „Mimikerkennung" wäre ein treffenderer Begriff, denn es kann – im psychologischen Sinne – nicht wirklich erkannt werden, welche Emotionen jemand hat, sondern nur, welchen *Gesichtsausdruck* jemand zeigt bzw. wie er oder sie sich verhält. Über Gestik und Mimik soll außerdem ermittelt werden, wer in einer Runde von Personen spricht - sogenannte *Sprechererkennung*.[148] Neben Mimik und Gestik wird auch versucht, die Körpersprache zu analysieren. Ziel ist es, Aggressivität zu messen (Abb. 4) oder medizinische Notfälle zu erkennen.

Im Forschungsprojekt *ADIS*[149] werden für die Interpretation der detektierten Bewegungen Situationen und Gestiken in typische Referenzfäl-

145 **Dao** et al., „Abandoned Object's Owner Detection: A Case Study of Hybrid Mobile-Fixed Video Surveillance System", S. 405.

146 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 341.

147 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications", S. 10.

148 **Cristani** et al., „Look at Who's Talking: Voice Activity Detection by Automated Gesture Analysis".

149 ADIS steht für „automatische Detektion interventionsbedürftiger Gefahrensituationen". Das Projekt wird vom Bundesministerium für Bildung und Forschung (BMBF) finanziert.

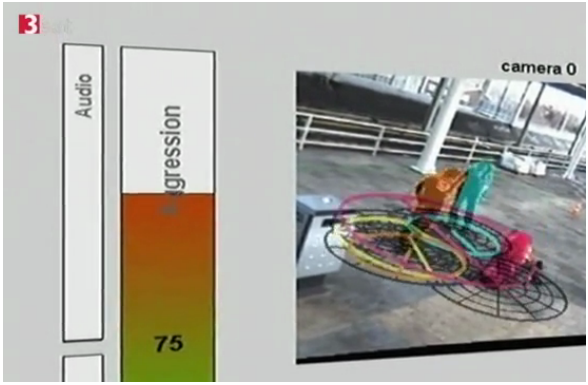


Abb. 4: Anhand der Bewegungen von Gliedmaßen soll ein Aggressionslevel gemessen werden.

le eingeteilt, die dann mit „realen Situationen" abgeglichen werden sollen.[150] Eine ähnliche Aufgabe konnte von Menschen nicht erfolgreich erfüllt werden. Im Rahmen des *screening passengers by observation techniques (SPOT) programme* sollte das Sicherheitspersonal Personen erkennen, die ein potentielles Sicherheitsrisiko auf Flughäfen darstellen. Dazu sollten sie nach Gesichtsausdrücken und Körpersprache Ausschau halten, die auf die Möglichkeit hinwiesen, dass eine Person in irgendeine Form von Betrug verwickelt war und befürchtete entdeckt zu werden. Das Sicherheitspersonal sollte sich dabei auf Verhalten konzentrieren, dass von etabliertem Normalverhalten abwich und auf Stress, Angst oder Betrug hätte hindeuten sollte. Das Programm wies hohe Raten von Fehlalarmen und verpassten Gefahren auf und wurde als nicht erfolgreich angesehen. **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 12 ff.

150 **Röbke**, *Automatisierte Detektion interventionsbedürftiger Situationen durch Klassifizierung visueller Muster (ADIS).*

3.2.4 Personenidentifizierung

Ein ganz eigenes Forschungsgebiet ist die Identifizierung von Personen anhand biometrischer Merkmale. Für Videoüberwachung sind besonders Gesichtserkennung und Identifizierung am Gang (*human gait recognition*) von Interesse. Die biometrischen Eigenschaften Gang und Gesicht eignen sich für Identifizierung besonders, da sie im Gegensatz etwa zum Fingerabdruck kein Mitwirken und keine Einwilligung der Personen erfordert. Während die Erkennungsraten von Gesichtserkennung unter Laborbedingungen extrem hoch sind und nahe an hundertprozentige Erkennung heranreichen, haben die Verfahren für den Einsatz im öffentlichen Raum noch Probleme mit mangelnder Bildqualität und unvorteilhafter Beleuchtung. Um die Robustheit zu erhöhen, wird die Gesichtserkennung mit der Erkennung am Gang kombiniert.[151] Bei *human gait recognition* werden Charakteristika menschlicher Bewegung und Proportionen ausgewertet. In der Theorie reichen bereits die Gelenkwinkel eines Menschen zur Identifizierung aus.[152] Bei Videoüberwachung ist Gangart als biometrisches Merkmal gegenüber anderen vorteilhaft, da sie nicht nur ohne Kooperation vonstatten geht, sondern die benötigten Bildinformationen auch aus der Ferne, ohne Bedarf hoher Bildauflösungen erhoben werden können.[153] Es gibt jedoch starke intrapersonelle Unterschiede der Bewegung, die durch äußere Umstände wie Schuhwerk, Kleidung, Stimmung oder Wegoberfläche beeinflusst werden und eine Identifizierung erschweren. Mit Hilfe statistischer Methoden soll am Gang einer Person außerdem auch das Geschlecht erkannt werden.[154]

151 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 343.
152 Ebd., S. 343.
153 **Kreissl**, „The effectiveness of surveillance in preventing and detecting crime and terrorism", S. 186.
154 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 343.

Gesichtserkennung bei Videoüberwachung kam und kommt bereits zum Einsatz. Im Londoner Vorort Newham werden gefilmte Gesichter mit Fahndungsfotos abgeglichen. Wenn eine der 300 Kameras eine 80-prozentige Übereinstimmung feststellt, wird die Polizei alarmiert.[155] Im Mainzer Hauptbahnhof[156] und 2008 auch in Berlin Kreuzberg am U-Bahnhof Kottbuser Tor[157] wurden Tests mit Gesichtserkennung durchgeführt. Ein viermonatiger Probelauf vom Bundeskriminalamt im Mainzer Hauptbahnhof war „nicht viel versprechend". Die Erkennungsrate sank nachts bei Kunstlicht auf zehn bis zwanzig Prozent.[158]

3.3 Vernetzung, Kameras und Sensoren

Bei der Entwicklung von Kameras wurde nicht nur die Bildqualität verbessert. Immer mehr Kameras haben integrierte Rechenleistung und können so Teile der Bildverarbeitung übernehmen und eigenständiges Teil eines Netzwerks werden. Kameras können auf diese Weise als Sammelstellen für Daten weiterer Sensoren fungieren.

3.3.1 Kameratypen

Neben Kameras mit CCD-Sensoren sind Nachtsichtgeräte und Wärmebildkameras die am meisten genutzten Geräte auf dem Videoüberwachungsmarkt.[159] In *HDRC-Kameras* (engl. *High Dynamic Range CMOS*) werden Bauelemente verwendet, die die Helligkeit nicht linear, sondern logarithmisch erfassen, so dass sie im Gegensatz zu CCD-Sensoren auch

155 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 69.
156 **Monroy**, *Allround-System für europäische Homeland Security*.
157 **Krempl**, *Berlin will Videoüberwachung mit biometrischer Gesichtserkennung testen*.
158 Ebd.
159 **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications", S. 279.

unter extremen Lichtbedingungen eingesetzt werden können.[160] Schon seit Längerem im Einsatz sind *Pan-tilt-zoom cameras* (PTZ). Sie können über Gelenke geschwenkt (*pan*) und geneigt (*tilt*) werden und verfügen über optischen Zoom. Oftmals werden inzwischen jedoch statt der teuren, mechanischen PTZ-Kameras eine Vielzahl billiger, statischer Weitwinkel- oder Domekameras so installiert, dass sich Bildbereiche überlappen.[161] Bei *Domekameras* sitzt die eigentliche Kamera hinter einer halbkugelförmigen meist getönten Blende. Viele Modelle können vollständig um die eigene Achse gedreht werden und äußerst schnelle Positionierungszeiten erreichen.[162] Von außen ist auf Grund der Tönung meist nicht ersichtlich, auf welchen Bereich die Kamera ausgerichtet ist. Von der Polizei werden mittlerweile auch mit Kameras ausgestattete Drohnen eingesetzt.[163] Auch die Deutsche Bahn kündigte vor kurzem an, mit Infrarotkameras ausgestattete Drohnen zur „Graffiti-Bekämpfung" einsetzen zu wollen.[164] Derartige Drohnen können über Geopositionierung wie GPS eigenständig Flugstrecken bestimmen und so automatisch vorher definierte Gebiete beobachten oder Objekte und Personen verfolgen.

3.3.2 Netzwerkkameras und Kameras mit Rechenleistung

Mit der Einführung netzwerkfähiger IP-Kameras ging der Trend außerdem hin zur Ausstattung der Kameras mit Rechenleistung und Speicher zur Vorverarbeitung des Bildmaterials dicht am Ort der Erhebung. In den meisten Fällen übernehmen die sogenannten *smart cameras* die *Bewegungsdetektion* und die *Objektverfolgung*, also die unteren Schichten der oben beschriebenen Hierarchie zur Bildanalyse und Verhaltenserken-

160 **Erhardt**, *Einführung in die Digitale Bildverarbeitung : Grundlagen, Systeme und Anwendungen*, S. 45.
161 **Moncrieff**, **Venkatesh** ; **West**, „Dynamic Privacy in Public Surveillance", S. 25.
162 **Scholz**, „§ 6b 2.2", S. 672.
163 **Monroy**, *Mehr Polizeidrohnen im Anflug*.
164 **Futurezone.at (Hrsg.):** *Deutsche Bahn will Drohnen gegen Sprayer*.

nung. Technisch gesehen sind die Verarbeitungsmöglichkeiten jedoch nur durch die Leistungsfähigkeiten der Prozessoren beschränkt, die ihre Aufgabe in der Praxis in Echtzeit ausführen müssen. Durch die Vorverarbeitung müssen nicht mehr alle Videodaten über das Netzwerk geschickt werden, sondern können die wesentlich kompakteren Ergebnisse, z. B. Trajektorienstücke von Objekten mit IDs, bandbreiteschonend übermittelt werden. Durch den integrierten Speicher können Videodaten vorgehalten und bei Bedarf übermittelt werden. Durch die dezentrale Verarbeitung und den reduzierten Datenverkehr skalieren Kameranetzwerke besser.

3.3.3 Integration von Sensoren

Zur Extraktion von Daten aus dem überwachten Raum werden nicht nur Kameras genutzt. Immer mehr finden auch andere Sensoren in der Forschung und bei der Planung von Überwachungsmaßnahmen Beachtung.[165] *Massively Integrated Multiple Sensor Installations* (MIMSI) wird als neuer vielversprechender Ansatz für Überwachungstechnik angesehen. Neben visuellen sollen beispielsweise auch akustische, chemische und elektromagnetische Sensoren zum Einsatz kommen.

Akustische Sensoren und Akustikanalyse

Auf dem 16. Berliner Polizeikongress wurde erklärt, in Kameras integrierte Mikrofone könnten es ermöglichen, unvermutete Geräuschpegel und verdächtige Gespräche einzufangen.[166] Auch beim EU Forschungsprojekt *INDECT* heißt es, es solle Ton aufgenommen werden, der nach charakteristischen Merkmalen abnormaler Situationen durchsucht werden sol-

165 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy", S. 324.
166 **Behörden Spiegel (Hrsg.):** *Mehr als nur Kamera-Überwachung.*

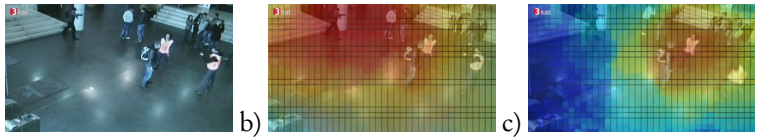


Abb. 5: Visualisierung von Schall im Kamerabild: a) Originalbild, b) Schuss und c) Schrei.

le.[167] Richtmikrofone können Gespräche aus etwa 100 Metern Entfernung erfassen[168] und mit Matrizen von Mikrofonen ist auch räumliches Hören möglich. So soll die Quelle eines auffälligen Geräuschs verortet und z. B. im Videobild eingeblendet werden (Abb. 5).

In Schottland wurden bereits Test zur Akustikanalyse durchgeführt.[169] Eine niederländische Firma stellte ihr System für zwei Wochen in einer belebten Straße in Glasgow auf. Das System belauschte oder zeichnete keine Gespräche auf, sondern wertete aus, *wie* etwas gesagt wurde. Das System könne zwischen aggressivem Ton einer Person und anderen lauten Geräuschen wie vorbeifahrende LKW unterscheiden.

Weitere Sensortypen

Auf Flughäfen werden bereits chemische Sensoren eingesetzt. Das technische Riechen nach explosiven Stoffen ist momentan auf einer recht fortgeschrittenen Stufe.[170]

Erwähnung finden in der Fachliteratur außerdem Diebstahlsensoren, Glasbruch-, Feuer-, Temperatur- und Feuchtigkeitsensoren sowie Lichtschranken und einfache Schalter.[171] An anderen Stellen werden Radar-, Ultra-

167 **Dziech**, „INDECT“, Folie 12.
168 **Moncrieff**, **Venkatesh** ; **West**, „Dynamic Privacy in Public Surveillance“, S. 25.
169 **Cannataci**, „Squaring the Circle of Smart Surveillance and Privacy“, S. 324.
170 **Macnish**, „Unblinking eyes : the ethics of automating surveillance“, S. 21.
171 **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces“, S. 1 ff.

schall- und Infrarotsensoren angesprochen.[172] Technisch ebenfalls in Frage kommen Lesegeräte für RFID-Chips.
Zur Auswertung von Kaufverhalten in Läden sollen in Kombination mit Kameras außerdem Sensoren für drahtlose Netzwerke genutzt werden.[173] Über Funksignale sollen Wege der Personen verfolgt und mit den Kamerabildern abgeglichen werden. Über eindeutige Codes, die von Mobiltelefonen beim Suchen nach Netzwerken gesendet werden, können Kunden (bzw. deren Mobiltelefone) auch bei einem späteren Besuch wiedererkannt werden.

Schnittstellen für beliebige Sensoren

Mitunter wird in der Literatur nicht konkretisiert, welche Art von Sensoren integriert werden sollen. In Veröffentlichungen des *INDECT*-Projektes wird neben der kombinierten Auswertung von Bild und Ton auch die Auswertung von „alphanumerischen Daten" erwähnt.[174] Es wird angestrebt, die Systeme variabel zu gestalten.[175] Es soll daher Hardware zum Einsatz kommen, die Standardschnittstellen zu den unterschiedlichsten Sensortypen zu Verfügung stellen. Diese Module sollen sich mitunter kabellos mit den Sensoren verbinden und die Werte über das Netzwerk dem Überwachungssystem zur Verfügung stellen.[176]

3.4 Weitere Datenquellen und ihre Bedeutung

Zur Überwachung können nicht nur Daten ausgewertet werden, die praktisch unmittelbar vor oder während der Verarbeitung erhoben wurden, sondern auch Daten und Informationen herangezogen werden, deren Ge-

172 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 347.
173 **Clifford** ; **Hardy**, *Attention, Shoppers: Store Is Tracking Your Cell*.
174 **Dziech**, „INDECT", Folie 11.
175 **Ruegg**, **November** ; **Klauser**, „CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples", S. 424.
176 **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces", S. 3.

winnung weiter zurück liegt. Daher schließt sich der Logik des Informationsflusses folgend nach Erhebung und Analyse nun die Datenspeicherung und der Zugriff auf Daten für spätere Analysen an. Nach der strukturellen Logik eines Überwachungssystems, an der sich dieses Kapitel ebenfalls orientiert, folgt die Beschreibung der Komponente, die für Speicherung und Zugriff zuständig ist.
Weiterhin mit dem Ziel, die Möglichkeiten der automatisierten Videoüberwachung darzulegen, soll unterschieden werden zwischen Datenspeichern, die Daten beherbergen, die vom Überwachungssystem selbst erhoben wurden – sie seien *interne Datenspeicher* genannt – und Datenquellen, die andere Daten bereitstellen – sie seien *externe Datenquellen* genannt. Letztere werden in 3.4.2 beschrieben.

3.4.1 Datenspeicher und rückwärts gerichtete Überwachung

In internen Datenspeichern können „Rohdaten" und Analyseergebnisse verschiedener Abstraktionsebenen abgelegt werden. Mit Rohdaten ist nicht gemeint, dass die Daten im ursprünglichen Format vorliegen sondern, dass sie nicht Ergebnis einer abgeschlossenen z. B. semantischen Analyse sind. Auch Videobilder, die in komprimierter Form gespeichert vorliegen, sollen daher als Rohdaten verstanden werden. Sicherlich ist der Übergang hier fließend und hängt zusätzlich von der Abstraktionsebene ab, um die es geht. Die Unterscheidung soll hier deutlich machen, dass einerseits jegliche Analyse, die sonst nur in Echtzeit möglich wäre, auch im Nachhinein noch durchgeführt werden kann und andererseits jede ermittelte Information auch später noch zur Verfügung stehen kann. Liegen Rohdaten vor, muss im Moment der Erhebung nicht bekannt sein, nach welchen Kriterien die Daten gefiltert, durchsucht und analysiert werden sollen. Dies erlaubt eine *zeitlich rückwärts gerichtete Videoüberwachung*. Während bei *manueller Videoüberwachung* alte Daten vorwiegend zur

Aufklärung von Geschehnissen genutzt werden konnten, können diese bei automatisierter Videoüberwachung auch im Echtzeitbetrieb herangezogen werden. Dabei ist – zumindest technisch – nur durch Speicher- und Rechenkapazität begrenzt, wie weit diese Daten in die Vergangenheit zurückreichen. Das vom Bundesministerium für Bildung und Forschung finanzierte Projekt *APFel*[177] verfolgt z. B. das Ziel Laufwege verdächtiger Personen rückwirkend zu ermitteln. Ein anderes Beispiel ist die zeitlich rückwärts gerichtete Fahndung nach Personen mit Gesichtserkennungsalgorithmen.[178]

An dieser Stelle kann nicht weiter auf Techniken zu Erstellung und Abfrage von Datenbanken eingegangen werden. Erwähnt sei jedoch, dass beispielsweise im Rahmen des *INDECT*-Projektes daran geforscht wird, Rohdaten und Analyseergebnisse so abzulegen, dass diese auch effizient und umfangreich im Nachhinein durchsucht werden können. Dazu wird das Videomaterial mit Metainformationen indiziert.[179]

3.4.2 Externe Datenquellen

Neben internen können auch externe Datenquellen zur Überwachung ausgewertet werden. Das können bestehende Datenbanken wie Melderegister, KFZ-Register, Fahndungsdatenbanken, Fluggastdaten, Biometriedatenbanken etc. sein, von denen konkrete Datensätze angefordert werden. Die Anbindung an eine KFZ- und Gesichtsdatenbanken wird eindrucksvoll in einem Präsentationsvideo des *INDECT*-Projektes vorgestellt (Abb. 6).

Auch Quellen, die mit dem Internet verbunden sind, können nach zusätzlichen Informationen durchsucht werden. Einige Informationen ste-

177 *APFel* steht für „Analyse von Personenbewegungen an Flughäfen mittels zeitlich rückwärts- und vorwärts gerichteter Videodatenströme".

178 Vgl. **Krempl**, *Berlin will Videoüberwachung mit biometrischer Gesichtserkennung testen*.

179 **Dziech**, „INDECT", Folie 10.

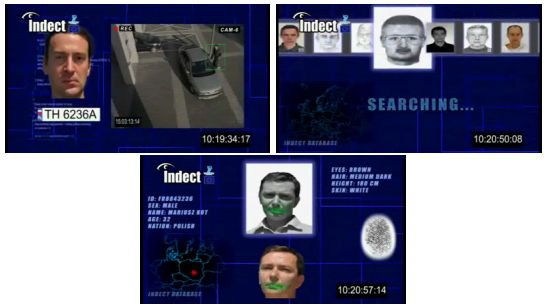


Abb. 6: Verknüpfung mit KFZ- und Gesichtsdatenbank in einem Präsentationsvideo des INDECT-Projekts.

hen frei zur Verfügung. So sind beispielsweise Strafregister aller 50 US-Bundesstaaten über die Internetplattform *criminalsearches.com* zugänglich. Die Daten stehen in aufgearbeiteter Darstellung und sogar inklusive biometrischer Porträtfotos zur Verfügung. So könnten mit Gesichtserkennung auch bei privat betriebenen Überwachungsmaßnahmen nach vermeintlich Kriminellen gesucht werden.

Neben solchen strukturierten Daten können auch mehr oder weniger unstrukturierte Daten aus dem Internet gewonnen und durch deren Verarbeitung weitere Informationen gewonnen werden. Ziel des Workpackage 4 des *INDECT*-Projektes ist es, mit Sprachverarbeitung (*Natural Language Processing*) und Maschinenlernen, Beziehungen zwischen Personen und Organisationen über Websites und soziale Netzwerke zu erheben.[180] Über Datamining, also dem Suchen nach Strukturen in unstrukturierten großen Datenmengen sind eine ganze Reihe von Analysen auch von Kommunikationsdiensten wie Twitter oder anderen Internetplattformen

180 **Klapaftis**, **Manandhar** ; **Pandey**, *XML Data Corpus: Report on methodology for collection, Deliverable name cleaning and unified representation of large textual data from various sources: news reports, weblogs, chat*, S. 7.

wie Blogs und Foren möglich. Auf diese Analysemöglichkeiten kann im Rahmen dieser Arbeit nicht weiter eingegangen werden.
In Anbetracht der Enthüllungen[181] durch den ehemaligen Mitarbeiter des US-amerikanischen Geheimdienstes NSA, Edward Snowden, über die weltweite, ungerichtete Überwachung der gesamten digitalen Datenverarbeitung und Telekommunikation duch US-amerikanischen und britischen Geheimdienste muss mit weit mehr als nur den oben beschriebenen Datenquellen gerechnet werden. Die Reaktionen der Regierungen lassen außerdem vermuten, dass von derartiger Praxis auch in Zukunft nicht abgewichen wird. Es muss vielmehr erwartet werden, dass nicht nur externe Daten bei automatisierter Videoüberwachung genutzt werden, sondern umgekehrt sowohl automatisierte als auch analoge Systeme selbst als Informationsquelle dienen, die allumfassende Überwachung des Digitalen um Informationen über Geschehnisse im *öffentlichen Raum* zu ergänzen. Tatsächlich war schon lange vor den Veröffentlichungen durch Snowden bekannt, dass die US-Regierung bei Google angefragt hatte, ob es technisch möglich sei, alle privaten Kameras so zu vernetzen, dass staatliche Behörden *jederzeit* auf sie zugreifen könnten.[182] Es muss also damit gerechnet werden, dass Videoüberwachungssysteme (manuelle oder automatisierte) auch Schnittstellen nach außen für den Zugriff auf Videobilder und internen Speicher zur Verfügung stellen werden.

3.5 Darstellungsansätze für die Mensch-System-Interaktion

Im Folgenden wird eine Auswahl möglicher Darstellung von Informationen für die OperateurInnen vorgestellt. Die Betrachtung ist relevant, da mit der Gestaltung der Interaktion zwischen Mensch und Überwa-

181 **Poitras**, *PRISM Whistleblower: Hong Kong*.
182 **Dix**, „Datenschutz und Informationsfreiheit : Bericht 2010“, S. 10.

chungssystem die Rolle und Aufgabe der OperateurInnen festgelegt wird und maßgeblich geprägt wird, auf welche Art und Weise diese erfüllt werden. Die Darstellungsweise von Informationen ist von besonderer Bedeutung für die später folgende Betrachtung der Verantwortung der OperateurInnen, da sie Entscheidungen auf Basis von Informationen fällen sollen, die vom System gefiltert und interpretiert dargestellt werden. Bisher wurden nur wenige Publikationen zur Visualisierung von Analyseergebnissen automatisierter Videoüberwachung veröffentlicht.

Die folgenden Ansätze zeigen jedoch bereits, dass zukünftige Systeme nicht mehr auf die Anzeige der Videobilder einzelner Kameras auf einer Matrix von Monitoren (Abb. 2) beschränkt sind. Zwei grundsätzliche Ansätze sind die Projektion der Videobilder in virtuellen 3D-Welten, in die weitere Informationen eingeblendet werden können, sowie das Anzeigen textueller Beschreibung von Geschehnissen in natürlicher Sprache.

Projektion der Videobilder in 3D-Welten

Eugster und Nebiker beschreiben, wie Videobilder einer Flugdrohne in Echtzeit in die Darstellung einer dreidimensionalen Umgebung integriert werden können (Abb. 7).[183]

Vergleichbare Verfahren werden in verschiedenen Veröffentlichungen beschrieben, die z. B. von Josef Scheuer thematisiert werden.[184] Philip DeCamp et al. erweitern das Prinzip für ein interaktives System namens *HouseFly*.[185] Es werden Bilder mehrerer, in verschiedenen Räumen eines Gebäudes verteilter Deckenkameras in eine Polygon-Repräsentation des Gebäudes projiziert und zu einer einzigen Ansicht verschmolzen (Abb. 8). Die Firma *Coherent Synchro* präsentiert eine kommerzielle Software,

183 **Eugster** ; **Nebiker**, „UAV-Based augmented monitoring : real-time georeferencing and integration of video imagery with virtual globes".

184 **Scheuer**, „Supporting Video Surveillance by Computer Graphics".

185 **DeCamp** et al., „An immersive system for browsing and visualizing surveillance video".

Abb. 7: Integration der Videobilder einer Flugdrohne in eine 3D-Umgebung.

die das freie Manövrieren in Echtzeit in einer 3D-Welt ermöglicht, in die die Videobilder ebenfalls perspektivisch hineinprojiziert werden.[186]
In diese Ansichten können z. B. Trajektorien von Laufwegen (Abb. 10), Transkripte oder zusammenfassende Wortwolken von erfassten Gesprächen (Abb. 9) oder etwa Fortschrittsanzeigen von Transaktionen einer Person an einem Verkaufsschalter (Abb. 11) an jeweiliger Stelle eingeblendet werden.
3D-Welten (bisher ohne das Projizieren von Videos) kommen bereits in kommerziellen Videoüberwachungssystemen zum Einsatz. Ein niederländischer Anbieter bewirbt das System *BeWare* bei dem Informationen in Videobilder, Kartenmaterial und 3D-Ansichten eingeblendet werden können. Hier wird beispielsweise das prozentuale Risikolevel für Gebäude, Gegenstände und Personen als farbiges Balkendiagram und Popupfenster dargestellt oder Laufwege und Aufenthaltsort von Personen in der Karte oder der 3D-Ansicht angezeigt (Abb. 12). Zusätzlich werden

186 **Coherent Synchro**, *Coherent Synchro 3D Visualization Platform*.

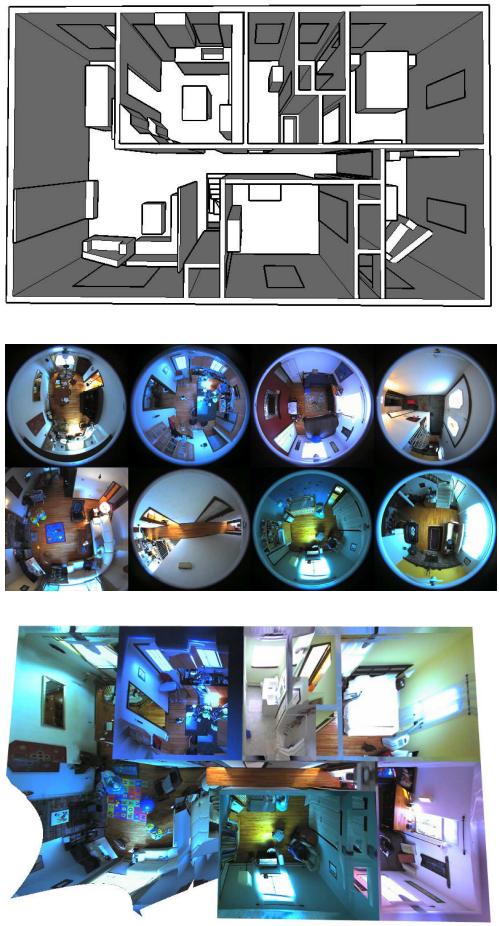


Abb. 8: Synthetisierung einer Gesamtansicht aus 3D-Modell und Kamerabildern in *HouseFly*.

Abb. 9: Lokalisiertes Einblenden von Gesprächstranskripten, die zu Begriffswolken zusammengefasst wurden.

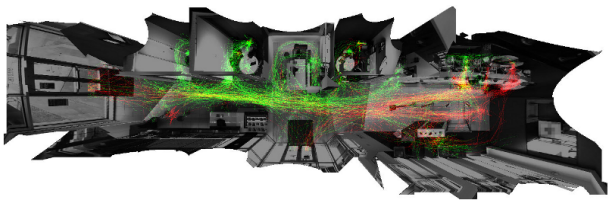


Abb. 10: Einblenden von Trajektorien in 3D-Ansicht.

Abb. 11: Fortschrittsanzeige der Interaktionsprozesse von Menschen an Schaltern.

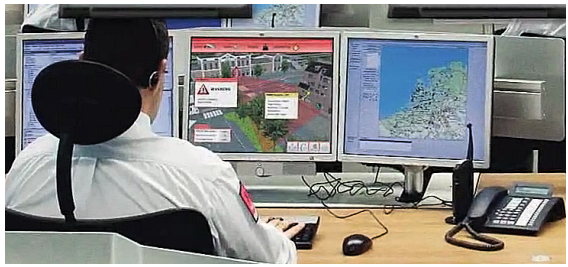


Abb. 12: Arbeitsplatz am Überwachungssystem *BeWare* mit 3D-Ansicht und Risikoanzeige.

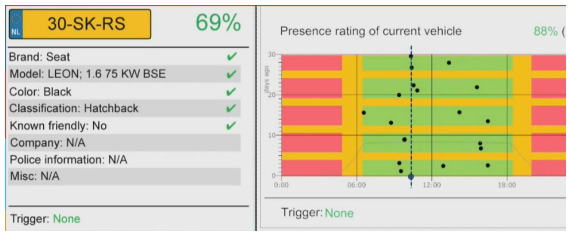


Abb. 13: Zeitliche Darstellung der Sichtungen eines konkreten Fahrzeuges.

Objekte im Videobild durch Umrandung hervorgehoben und mit zuvor gesammelten Informationen annotiert. Dazu gehört z. B. auch die chronologische Darstellung der registrierten Aufenthalte eines Fahrzeuges im überwachten Gebiet (Abb. 13).

Textuelle Beschreibung

Eines der Ziele von Videoüberwachung nach Weiming Hu et al. ist es, das Verhalten von Objekten in kurzer, klarer und natürlicher Sprache wiederzugeben, da dies auch der geeignetste Weg sei, wie Menschen untereinander kommunizieren würden.[187] Nach Robertson ist Ziel nicht nur die Be-

187 **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 342.

schreibung dessen was passiert, sondern vornehmlich die Erklärung der Interaktionen, die stattfinden.[188] Mit einem regelbasierten System sollen auf Basis von Trajektorien beispielsweise Aussagen wie „Person A läuft vom Gehweg auf die Straße, um Person B auszuweichen" getroffen werden.[189]

188 **Robertson**, **Reid** ; **Brady**, „Automatic human behaviour recognition and explanation for CCTV video surveillance", S. 1.

189 Ebd., S. 13.

3.6 Versuche *privacy*-fördernder Techniken

Die in 2.3.2 dargestellten Datenschutz- und Diskriminierungsproblematiken bei Videoüberwachung werden seit einigen Jahren auch bei der Forschung an Techniken zur Automatisierung thematisiert. Die computerisierte Verarbeitbarkeit der Bilder wird als Chance betrachtet, gegenüber *manueller Videoüberwachung* datenschutzfreundlichere Techniken (*Privacy Enhancing Technologies (PET)*[190]) zu entwickeln. An dieser Stelle sollen die technischen Ansätze zur Förderung von *privacy* zunächst vorgestellt werden. In 5.1.2 werden sie diskutiert.

Ziel der nun vorgestellten Techniken ist es, den Zugang zu Informationen zu unterbinden, die in die *privacy* der Individuen eingreift.[191] In Kapitel 2.3.2 wurde *privacy* als Konzept vorgestellt, das von Person zu Person und Gruppe zu Gruppe unterschiedlich ausgeprägt ist. Genau zu spezifizieren, welche Informationen bei Videoüberwachung sensitiv sind, wird daher als schwierig angesehen.[192]

Privacy als Optimierungsproblem

Moncrieff et al. sehen *privacy* aus Sicht der Entwicklung von Überwachungstechnik daher als ein Optimierungsproblem bezüglich des Informationsflusses an. Nähme man an, dass Menschen generell so wenig wie möglich Informationen über sich preisgeben wollen würden, so sollte ein Überwachungssystem die maximale Datenmenge liefern, die die Überwachenden benötigen, um ihr Ziel zu erreichen, während dabei der Eingriff in die *privacy* der Überwachten minimiert werde.[193]

Ein Optimum wäre zu suchen zwischen vollständigem Ausblenden aller Personen im Bild, wodurch das Videomaterial nutzlos wäre, und einem maximalen Level von Informiertheit, bei der Personen im überwachten

190 **Deutscher Bundestag (Hrsg.):** *Drucksache 17/3940*, S. 9.
191 **Senior**, *Protecting Privacy in Video Surveillance*, S. 37.
192 Ebd.
193 **Moncrieff**, **Venkatesh** ; **West**, „Dynamic Privacy in Public Surveillance", S. 25.

Bereich z. B. RFIDs zur lückenlosen individuellen Beobachtung bei sich tragen müssen.[194]

Ansätze

Es gibt verschiedene Ansätze, mit denen versucht wird, den Eingriff in die *privacy* bei Videoüberwachung zu verringern und gleichzeitig die Überwachungsaufgabe zu ermöglichen. Es kann eine Einschränkung der Beobachtbarkeit vorgenommen werden, Techniken zum Verstecken von Daten im Videomaterial (*datahiding*) eingesetzt werden und Konzepte zur Zugriffsbeschränkung durch Rechtevergabe angewendet werden. Außerdem wird generell angestrebt, den Bedarf an Auswertung der Videobilder durch Menschen zu minimieren, um die *privacy* gegenüber den OperateurInnen zu schützen und Problemen wie Missbrauch und Voyeurismus aus dem Weg zu gehen.[195]

3.6.1 Verhinderung oder Einschränkung der Bildaufnahme

Ein simpler Ansatz ist das Verhindern oder Einschränken der Beobachtung bestimmter Bereiche. Dies kann erreicht werden durch physische Einschränkungen wie Blenden[196] oder Einschränkung der Schwenkwinkel der Kameras oder per Software zum Beispiel durch Unterbinden des Zoomens oder des Schwenkens in vorher festgelegte Bereiche.[197] Derartige Möglichkeiten sind jedoch sehr begrenzt.

3.6.2 *Datahiding*

Eine weitere Möglichkeit ist das softwaretechnische Ausblenden bestimmter Bereiche der Bilder, genannt *datahiding*. *Datahiding* wird algorith-

194 Ebd., S. 25.
195 **Senior**, *Protecting Privacy in Video Surveillance*, S. 35.
196 **Senior** et al., „Enabling Video Privacy through Computer Vision", S. 52.
197 **Coudert** ; **Dumortier**, „Intelligent Video Surveillance Networks: Data Protection Challenges", S. 978.

misch in zwei Schritten umgesetzt. Zunächst müssen Bildbereiche mit privacy-relevanten Informationen (*regions of interest (ROI) lokalisiert* und anschließend *unkenntlich* gemacht werden. Andrew Senior et al. und Thomas Winkler stellen Kameras mit eigener Rechenleistung vor, die diese Aufgabe durchführen, noch bevor die Bilddaten dem restliche System zur Verfügung gestellt werden.[198]

Senior et al. beschränken sich bei der Auswahl der *ROI*, deren Sichtbarkeit die *privacy* einschränken würde, im allgemeinen auf Handlungen von Menschen und im Speziellen auf Merkmale die zur Identifizierung genutzt werden können.[199] Identifizierungsmerkmale werden in starke und schwache Merkmale eingeteilt. Starke Identifizierungsmerkmale sind Biometriedaten wie Gesicht und Gang, anhand derer eine genügend eindeutige Identifizierung möglich ist. Schwache Identifizierungsmerkmale hingegen sind Körpergröße, Schrittweite, oder Merkmale wie Farbe der Kleidung, die keine eindeutige Identifizierung zulassen.

Datahiding für Gesichter

Das vollständige Löschen von Gesichtern ist mitunter nicht erwünscht, da der Gesichtsausdruck der Beobachteten als Information für OperateurInnen verloren gehen würde. Zwei simple Ansätze zum mimikerhaltenden *datahiding* sind die Reduzierung der Auflösung der Gesichter (*Verpixelung*) oder das Herabsetzen der Schärfe durch Anwendung von Gaussfiltern (Abb. 15b und 15c).

Frederic Dufaux and Touradj Ebrahimi untersuchten die Effektivität dieser „naiven" *datahiding*-Verfahren mit Gesichtserkennungsalgorithmen und einer gängigen Gesichtsdatenbank zur Evaluation von Algorithmen.[200]

198 **Senior** et al., „Enabling Video Privacy through Computer Vision"; **Winkler**, „Vertrauenswürdige Videoüberwachung : Sichere intelligente Kameras mit Trusted Computing".

199 **Senior**, *Protecting Privacy in Video Surveillance*.

200 **Dufaux** ; **Ebrahimi**, „A framework for the validation of privacy protection solutions in video surveillance".

Das mit *datahiding* geschützte Gesicht wurde von einem Algorithmus mit der Gesichtsdatenbank abgeglichen. Der Algorithmus liefert aus der Datenbank eine Reihe nach Übereinstimmung sortierter Gesichter (Treffer). Nach einer Reihe von Versuchen für verschiedene Gesichter ergeben sich bestimmte Wahrscheinlichkeiten. Die Effektivität des *datahiding* kann dann mit der Erkennungsrate verschiedener *Ränge* angegeben werden. Dabei gibt die Erkennungsrate des Ranges n an, mit welcher Wahrscheinlichkeit die tatsächlich abgebildete Person unter den ersten n Treffern zu finden ist. Die Qualität des *datahiding* kann über die Ränge z. B. von eins bis fünfzig wie in Abb. 14 veranschaulicht werden. Bei einigen Gesichtserkennungsalgorithmen blieb die Erkennungsrate für Rang $eins$ nach Verpixelung oder Anwendung des Gaussfilters signifikant hoch, das heißt, Personen wurden trotz Verpixelung noch mit hoher Warhscheinlichkeit erkannt. Die beiden Verfahren wurden von Dufaux und Ebrahimi als untauglich eingestuft. Außerdem wurde darauf hingewiesen, dass auch bei stärkerer Verpixelung eine höhere Auflösung durch Integration der Trajektorien von Pixeln über mehrere Bilder wieder errechnet werden kann.

Zwei weitere Verfahren (Abb. 15d und 15e) ergeben nach ihrer Anwendung im Rang $eins$ wesentlich geringere Erkennungsraten von nahezu null Prozent. Sie werden daher für die Verhinderung automatisierter Identifizierung als tauglich angesehen. Nichtsdestotrotz steigen die Erkennungsraten von Rang 0 bis Rang 50 auf fast 10 Prozent (Abb. 14). Eine gewisse Korrelierbarkeit bleibt also auch hier bestehen.

Es existieren auch weitere Ansätze, die Details von Gesichtern grundsätzlich erhalten, jedoch die Gesichter so verändern, dass Gesichtserkennungsalgorithmen diese nicht mehr verlässlich erkennen können sollen. Eine Variante davon ist, das Gesicht geometrisch hin zu einem Durch-

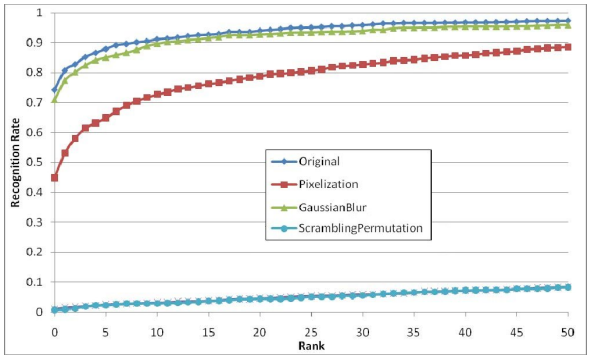


Abb. 14: Erkennungsraten nach unterschiedlichen Methoden der Unkenntlichmachung.

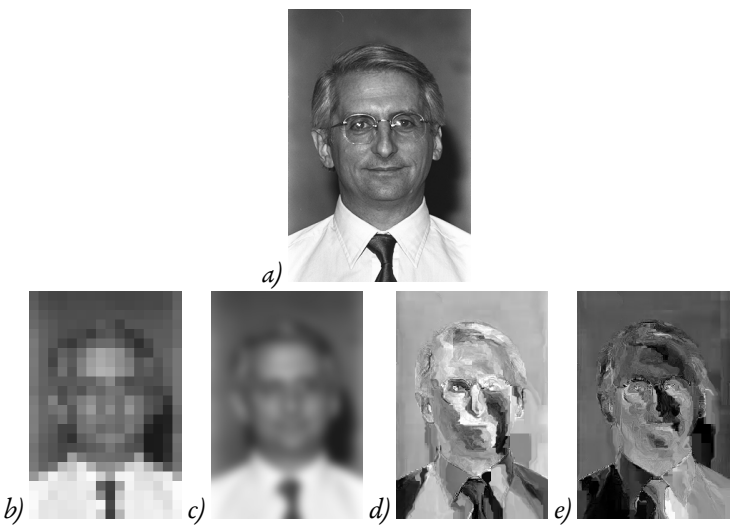


Abb. 15: Unkenntlichmachung von Original a) mittels Verpixelung *b)*, Gaussfilter *c)* und zwei weiteren Techniken.

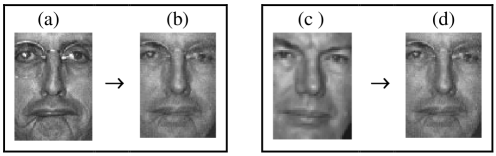


Abb. 16: Detailerhaltende Unkenntlichmachung von Gesichtern durch geometrische Annäherung an ein Durchschnittsgesicht (hier lediglich aus *a)* und *c)* gebildet).

schnittsgesicht zu verzerren.[201] Eine Garantie für Anonymität können auch diese Algorithmen nicht leisten.[202]

Weiterhin gibt es *reversible* Verfahren zum *datahiding*, bei denen die *ROI* mit einem geheimen Schlüssel wieder hergestellt werden können. Es wird nur ein Videostream erzeugt, übertragen und gespeichert, der die verschlüsselten *ROI* enthält. Nur diejenigen, die Zugriff auf den geheimen Schlüssels haben, können die *ROI* entschlüsselt betrachten.

Datahiding für den Körper

In einer Studie über das individuelle Verständnis von *privacy* wurde mit Fragebögen das Verhältnis zu Überwachung in Kommunen ausgewertet.[203] Man kam zu dem Ergebnis, dass je mehr man einen Überwachenden kenne, desto mehr vertraue man diesen und um so mehr würde man bereit sein von sich preiszugeben. Zum Beispiel könnte man damit einverstanden sein, dem freundlichen Hausmeister auf seinem Bildschirm beim Betreten des eigenen Hauses zu erscheinen. Die anonymen MitarbeiterInnen einer Sicherheitsfirma sollen jedoch nur sehen können, *dass* jemand sich im Eingangsbereich aufhält, jedoch nicht wer. Das Projekt *PriSurv* versuchte eine Stufung von *datahiding* umzusetzen. Für den Schutz

201 **Senior**, *Protecting Privacy in Video Surveillance*, S. 37.
202 **Newton**, **Sweeney** ; **Malin**, „Preserving Privacy by De-Identifying Face Images“, S. 24.
203 **Senior**, *Protecting Privacy in Video Surveillance*, S. 147.

unterschiedlicher Informationen müssen diese auf unterschiedliche Weise aus dem Bild entfernt werden. Unterschieden werden Informationen wie Existenz, Aufenthaltsort, Körpergröße, Umfang, Silhouette, Frisur, Kleidung, Gesichtsausdruck oder Körpersprache (Abb. 18). Diese sollen aus dem Bildern entfernt werden durch den Austausch der Person mit der Silhouette, einem proportionalen Quader, einer Linie, einem Punkt oder durch die vollständige Entfernung der Person aus dem Bild (Abb. 17). Welche Personen auf welche dieser Informationen zugreifen dürfen, kann mit Zugriffskontrolle geregelt werden.

3.6.3 Zugriffskontrolle

Senior et al. schlagen eine schichtweise Zugriffskontrolle vor.[204] Drei verschiedenen Benutzergruppen soll unterschiedlich umfangreicher Zugriff gestattet werden. Einfache Nutzer bekommen Zugriff zu statistischen Daten, privilegierte Nutzer können begrenzt individuelle Information einsehen und Vollzugsbehörden sollen Vollzugriff auf die rohen Videoinformationen und personenbezogenen Daten erhalten. Zur Umsetzung können konventionelle Methoden wie Verschlüsselung und Listen mit Zugangsbeschränkungen genutzt werden. Zum Verbergen bestimmter Informationen in den Videobildern können die oben beschriebenen *datahiding*-Methoden zur Anwendung kommen. Entweder liegen die Originaldaten der *ROI* separat zum Video verschlüsselt vor, oder sind mit reversiblen Methoden verschlüsselt in den Bilddaten selbst enthalten.

Sicherheit der Systeme

Auf Sicherheit der Systeme, ein Aspekt der für den Datenschutz ebenfalls von Bedeutung ist, zum Beispiel Maßnahmen gegen *leaking* und *tapping*, kann hier nicht weiter eingegangen werden. Im Zusammenhang mit der

204 **Senior** et al., „Enabling Video Privacy through Computer Vision".

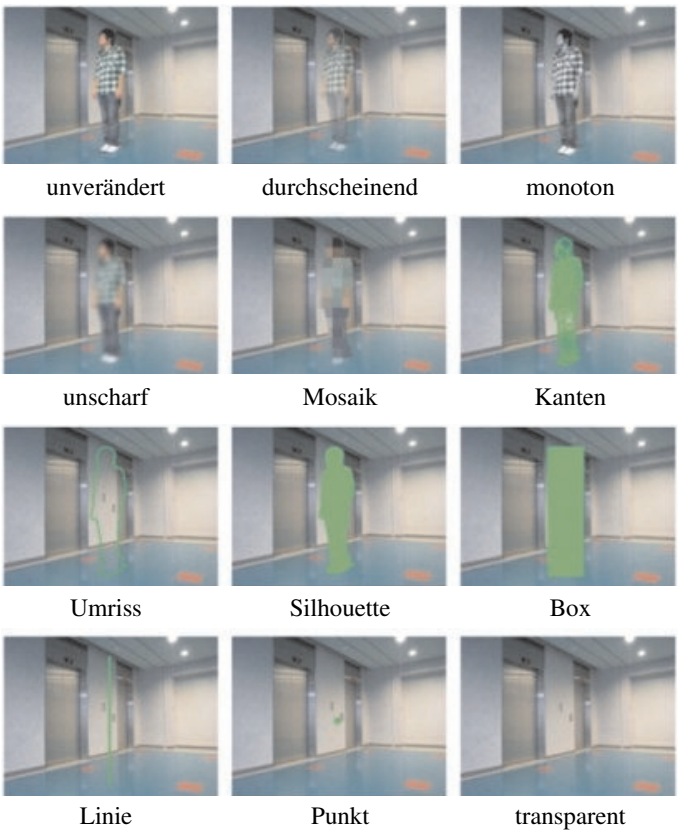


Abb. 17: unterschiedliche Stufen von *datahiding*.

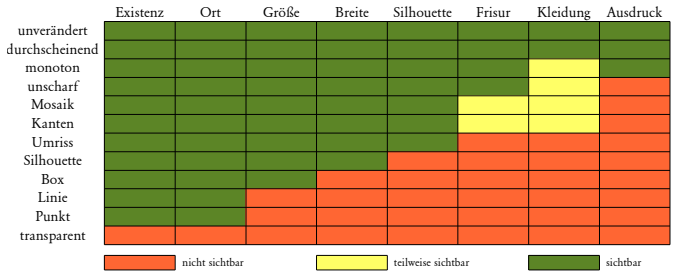


Abb. 18: Verschiedene Informationen werden ausgeblendet.

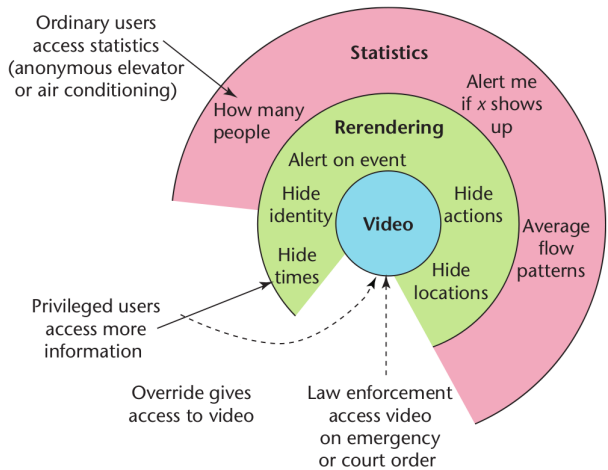


Abb. 19: Drei Benutzertypen erhalten unterschiedlich umfangreichen Zugriff auf Überwachungsinformationen.

Rechtevergabe soll nur erwähnt werden, dass mit Hilfe von Watermarking die Videodaten beim Anzeigen oder Exportieren je nach Betrachter oder Betrachterin mit Metadaten versehen werden sollen, die bei einem Leak Rückschlüsse auf verantwortliche Personen erlauben. Im Rahmen des *INDECT*-Projektes wird beispielsweise an Watermarking geforscht, das nicht trivial – etwa durch Recodierung – aus Videos entfernbar sein soll.[205]

Welche Daten erhoben werden und wer auf welche Informationen Zugriff hat; oder mit anderen Worten, wie stark Betroffene in ihrer *privacy* eingeschränkt werden, muss nicht allein von den Zugriffsrechten abhängen. Eingriffstiefe und Zugriffsumfang können auch dynamisch je nach Situation und Kontext angepasst werden:

3.6.4 Kontextuell-dynamische Eingriffstiefe in Grundrechte

Alexander Rossnagel, Monika Desoi und Gerrit Hornung schlagen ein *Drei-Stufen-Modell* vor, um den Eingriff auf informationelle Selbstbestimmung und weitere Grundrechte auf jeder Stufe in Abhängigkeit vom Grad der ermittelten Gefahr „so gering wie möglich zu halten".[206] Ein weiteres Ziel ist es, dem Überwachungspersonal bei seinen Entscheidungen ein handhabbares Schema vorzugeben, das im praktischen Einsatz die Unterscheidung zwischen rechtmäßigem und rechtswidrigem Eingriff ermöglicht.

Erste Stufe

In der *ersten Stufe* soll eine „allgemein beobachtende Überwachung" aller Personen stattfinden. Die Bewegungen der Personen sollen automatisiert mit definierten Mustern verglichen werden. Eine Identifizierung ist

205 **INDECT (Hrsg.):** *INDECT FAQ : Frequently Asked Questions*, S. 6.

206 **Roßnagel**, **Desoi** ; **Hornung**, „Gestufte Kontrolle bei Videoüberwachungsanlagen : Ein Drei-Stufen-Modell als Vorschlag zur grundrechtsschonenden Gestaltung".

nicht nötig, daher sollen Personen weitgehend mit *datahiding* unkenntlich gemacht werden. Möglichkeiten des Zoomens sind eingeschränkt und Daten werden nach kurzer Zeit gelöscht. Eventuell könne die Gesamtsituation vollautomatisch ohne menschliche Interaktion kontrolliert werden. Für den automatischen Übergang zur zweiten Stufe „ist es erforderlich aber auch ausreichend, wenn ein Kamerasystem oder der Beobachter ein Verhalten feststellen, das hinreichende Anhaltspunkte enthält, um bei verständiger und besonnener Lagebeurteilung eine Situation anzunehmen, die erfahrungsgemäß eine Gefahr verursacht."[207]

Zweite Stufe

In der *zweiten Stufe* („gezielte Personenüberwachung") wird der Beobachter auf die Situation aufmerksam gemacht. Das System verfolgt die markierte Person, wählt die beste Kameraperspektive aus und überprüft weiter das Verhalten. Der Beobachter kann zoomen, jedoch soll eine Aufnahme des Gesichtes und biometrischer Merkmale verhindert werden. Wird festgestellt, dass keine Gefahr vorliegt, sollen die Daten unmittelbar gelöscht werden. Der Übergang in die dritte Stufe kann ebenfalls automatisch oder manuell erfolgen:

> Für den Übergang in die dritte Stufe ist es erforderlich, aber auch ausreichend, wenn das Kamerasystem oder der Beobachter ein Verhalten feststellen, das bei verständiger und besonnener Lagebeurteilung eine konkrete unmittelbare Gefahr oder den konkreten Verdacht einer Straftat begründet. Eine Gefahr liegt vor, wenn eine Sachlage festgestellt wird, bei der im Einzelfall die hinreichende Wahrscheinlichkeit be-

207 **Roßnagel**, **Desoi** ; **Hornung**, „Gestufte Kontrolle bei Videoüberwachungsanlagen : Ein Drei-Stufen-Modell als Vorschlag zur grundrechtsschonenden Gestaltung", S. 695.

steht, dass in absehbarer Zeit ohne Eingreifen ein Schaden an Rechtsgütern verursacht wird.[208]

Dritte Stufe

Mit der *dritten Stufe* („Personenerkennung") werde in der Regel das Einschreiten von Sicherheitskräften erforderlich. Sie soll der Einsatzleitung und der Beweissicherung dienen. Nach wie vor soll keine automatische Identifizierung stattfinden, es sollen jedoch biometrische Daten erfasst werden, so dass eine Identifizierung später möglich ist. Mit Eintreten in die dritte Stufe können „Gefahrenabwehr und Strafverfolgung" erforderlich sein. Dazu sollen die gesammelten Daten genutzt werden. Die Daten können gelöscht werden, wenn „trotz der eindeutigen Hinweise doch weder eine Gefahr vorlag, noch eine Straftat verübt wurde".[209]

208 Ebd., S. 696.
209 Ebd., S. 700.

4 Entwurf eines technisch wahrscheinlichen Komplettsystems

In Kapitel 5 sollen gesellschaftliche Probleme automatisierter Videoüberwachung identifiziert werden. Es reicht nicht aus, die vorgestellten Techniken und Ansätze isoliert nach möglichen Probleme zu untersuchen, denn der Charakter automatisierter Videoüberwachung wird erst deutlich anhand der Architektur des Systems, der Eigenschaften der Komponenten in ihrem Zusammenspiel und auch der Ausgestaltung der Mensch-Technik-Interaktion.
Im Folgenden soll daher ein technisch wahrscheinliches System Ω anhand von vorhandenen Techniken und Systemen sowie technischen Trends, Anforderungen und Architekturvorschlägen[210] entworfen werden.

Entwurf mit vollem technischen Potential

Ω wird ganz bewusst, ohne rechtliche Einschränkungen zu berücksichtigen, mit dem vollen technischen Potential entworfen. Die rechtlichen Rahmenbedingungen für den Einsatz von Ω unterscheiden sich weltweit stark. In einigen Ländern werden Grundrechte unter Umständen wesentlich schlechter geschützt oder existieren gar nicht, so dass damit zu rechnen ist, das eingesetzt wird, was technisch machbar erscheint. Anhand momentaner Überwachungspraxis von Geheimdiensten wird deutlich, dass die Gesetze zum Schutz der Grundrechte sogar in Demokratien in großem Stil missachtet und alle technischen Möglichkeiten ausgenutzt werden. Dies spiegelt den zu beobachtenden Trend wider, dass auch in Demokratien nicht mehr der Schutz der Freiheit des Einzelnen durch Si-

210 Es wird sich vorwiegend an dem im März 2013 ausgelaufenen Projekt *CamInSens* und dessen Beschreibung von d'Angelo orientiert. Das Projekt *CamInSens* (Verteilte, vernetzte Kamerasysteme zur in situ-Erkennung personeninduzierter Gefahrensituationen) wurde im Programm „Forschung für die zivile Sicherheit" durch das *Bundesministerium für Bildung und Forschung* (BMBF) im Rahmen der High-Tech-Strategie gefördert.

cherheit, sondern die Sicherheit des Systems im Mittelpunkt steht.[211] Vor diesem Hintergrund erscheint es zwingend notwendig, Ω mit dem maximalen technischen Potential zu entwerfen, um ein möglichst umfangreiches Bild der Probleme herleiten zu können.

Aussagekraft des Entwurfs

Obwohl Videoüberwachung unterschiedlichste Einsatzgebiete und Dimensionen haben kann, ist nicht mit einer starken Variabilität der grundsätzlichen Funktionsweise der Systeme bzw. einzelner Komponenten zu rechnen. Außerdem vermitteln die Publikationen der Forschungsgebiete, dass in absehbarer Zeit nicht wesentlich von den bisher entwickelten, hier vorgestellten Herangehensweisen und Techniken abgewichen wird. Der Wunsch, die Systeme möglichst generisch, dynamisch, skalierbar und verknüpfbar zu gestalten, spricht außerdem dafür, dass sich Standards durchsetzen werden. Auch aus Gründen der kommerziellen Verwertung wird dies der Fall sein, denn es werden Module und Bauteile unterschiedlicher Firmen zum Einsatz kommen, bei denen bestimmte Schnittstellen vorgegeben werden sein. Der Entwurf eines generischen Systems Ω, das je nach Einsatzgebiet ausgestaltet wird, ist daher zur Identifizierung gesellschaftlicher Probleme zielführend.

4.1 Anforderungen an das System

Vorwärts und rückwärts gerichtete Überwachung

Ziel von Ω ist nicht nur das Erkennen von aktuell stattfindenden unerwünschten Geschehnissen. Es sollen auch in der Zukunft stattfindende Geschehnisse prognostiziert werden, so dass mit diesen rechtzeitig umgegangen werden kann. Dies soll vor allem über den Abgleich der Gescheh-

211 **Albrecht**, *Der Weg in die Sicherheitsgesellschaft : Auf der Suche nach staatskritischen Absolutheitsregeln*.

nisse mit automatisch erstellten Modellen von Normalität (*Whitelistansatz*) erreicht werden. Um die Korrektheit der Prognosen zu verbessern, werden jegliche zusätzlich verfügbaren Informationen bei der Analyse berücksichtigt.
Außerdem soll Ω auch eine zeitlich rückwärts gerichtete Überwachung ermöglichen, in dem Daten ausgewertet werden, deren Erhebung zeitlich zurückliegt. Dazu werden nicht nur fertige Analyseergebnisse sondern auch Rohdaten vorrätig gehalten. Die Suchkriterien und Modelle können so auch im Nachhinein festgelegt werden. Damit derartige Suchen von außen, z. B. auch über mehrere Systeme hinweg, veranlasst werden können, bietet Ω externe Schnittstellen für den Datenzugriff an.

Aufwandsminimierung

Da Ziel der Automatisierung unter anderem eine Kostenersparnis ist, soll sowohl der Konfigurationsaufwand des Überwachungssystems als auch der Personalaufwand während des Einsatzes möglichst gering sein. Dazu werden vorgefertigte Komponenten bereitgestellt, die je nach Einsatzgebiet zusammengestellt und über maschinelles Lernen größtenteils automatisch konfiguriert werden. Auch die Darstellung der Ergebnisse für OperateurInnen wird so gestaltet, dass nötige Fähigkeiten und Personalaufwand möglichst minimiert wird. Ebenfalls im Zusammenhang mit den Personalkosten steht das Ziel der Minimierung der Fehlalarme.

4.2 Struktur und Komponenten

Damit Ω auch für große Gebiete wie Bezirke oder ganze Städte skaliert, ist es möglichst dezentral organisiert. Noch Anfang dieses Jahrtausends wurden zentralistische Systeme mit einem einzigen „Threat Assessment Processor“ vorgeschlagen, der jegliche Analyseaufgaben übernahm und durch Programmierung an die jeweiligen Bedürfnisse angepasst werden

konnte.[212] Heutige Forschung strebt an, verteilte Systeme zu entwickeln, die großflächig Räume unterschiedlichster Art mit einer Vielzahl von Sensoren und Kameras abdecken sollen.[213] Für die integrierte Darstellung der Daten und der Ergebnisse beinhaltet jedoch auch Ω ein zentrales datenverarbeitendes Element (Abb. 20g) – die Dezentralität bezieht sich hauptsächlich auf das Kamera- und Sensornetzwerk (Abb. 20f) sowie die unteren Schichten der Bildanalyse. Weitere Bestandteile von Ω sind in Abb. 20 dargestellt und werden im Folgenden näher beschrieben.

Kameras

Bei der Installation von Ω wird eine große Anzahl von Kameras mit sich überlappenden Bildbereichen angestrebt. Dies zielt nicht nur auf eine umfangreichere Beobachtung ab, sondern hat vor allem auch technische Gründe. Eine ganze Reihe von Problemen des Bildverstehens wie räumliche Uneindeutigkeiten und Verdeckungen, die in realen Überwachungsszenarien durch hohes Personenaufkommen entstehen, lassen sich erst mit Bildern aus verschiedenen Perspektiven zuverlässig lösen.[214] Die Platzierung der Kameras hat dadurch einen enormen Einfluss auf die Leistung der Bildverarbeitung. Es gibt Verfahren, die Platzierung von Kameras aufgabenspezifisch optimieren.[215]. Dabei werden nicht nur Bildausschnitte sondern auch benötigte Bildauflösungen bestimmter Bereiche berücksichtigt. Zusätzlich zu den fest installierten Kameras können in Ω auch mit Kameras ausgestattete Flugdrohnen integriert werden, die bei Bedarf autonom vorgegebene Strecken abfliegen oder Objekte und Personen verfolgen können (Abb 20c). Die Videobilder können per Funk in Echtzeit an das restliche System übertragen werden.

212 **Thiel**, „Automatic CCTV surveillance-towards the VIRTUAL GUARD“, S. 8.
213 **Valera** ; **Velastin**, „Intelligent distributed surveillance systems: A review“, S. 202.
214 Vgl. **Hu** et al., „A survey on visual surveillance of object motion and behaviors“, S. 344.
215 **Bodor** et al., „Optimal Camera Placement for Automated Surveillance Tasks“.

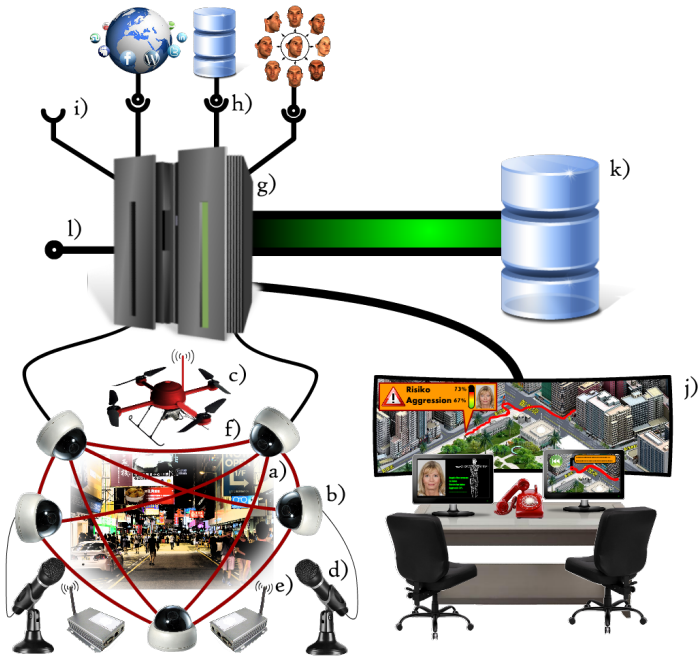


Abb. 20: Schematischer Aufbau von Ω.

a) Überwachter Raum.
b) Untereinander vernetzte Kameras mit Rechenleistung übernehmen einen Großteil des Bildverstehens.
c) Autonome Flugdrohnen übertragen Bilder an f).
d) Zusätzliche Sensoren z. B. Mikrofone sind über Kameras mit dem Netzwerk verbunden.
e) Weitere Sensoren sind per Funk am Netzwerk angeschlossen.
f) Skalierendes Netzwerk verbindet einzelne Komponenten.
g) Zentrale Datenverarbeitung integriert die Daten, wertet sie aus und stellt sie in j) dar.
h) Anbindung an externe Datenquellen wie Internetquellen oder Biometrie- und Personendatenbanken.
i) Generische Schnittstelle für weitere Datenquellen.
j) Darstellung der Kamerabilder, Analyseergebnisse und Informationen in virtueller 3D-Umgebung.
k) Speicher für Rohdaten und Analyseergebnisse zur rückwärts gerichteten Überwachung.
l) Schnittstellen für den Datenzugriff auf f) und k) von außen.

Sensoren

Zusätzlich zu den Kameras können unterschiedlichste Sensoren überall im überwachten Raum platziert sein (Abb. 20e). Diese können ihre Signale und Werte an die nahen Kameras oder andere am Netzwerk angeschlossene Module per Funk übertragen und von dort aus dem restlichen System zur Verfügung gestellt werden. Über generische Schnittstellen können beliebige Sensoren integriert werden. Mögliche Sensoren sind Mikrofone zur Auswertung und Verortung von Geräuschen und Sprache (Abb. 20d), RFID-Sensoren zum Auslesen von Ausweisdokumenten oder anderen RFID-Labels, WLAN-Sensoren zum Verfolgen und Identifizieren von Geräten wie Mobiltelefonen, chemische Sensoren zum Aufspüren von Substanzen wie Drogen oder Sprengstoff, und einfache Sensoren wie Glasbruchsensoren, Lichtschranken oder zur Messung von Temperaturen. Vom Gesamtsystem können die Sensordaten über Identifikationsnummern einem bestimmten Ort bzw. einem bestimmten Kontext zugeordnet werden und so in das Gesamtbild des überwachten Raumes integriert werden.

Externe Datenquellen und Schnittstelle nach außen

Zur Verarbeitung und Darstellung von Informationen, die nicht im Rahmen der Überwachung selbst erhoben wurden, stellt Ω Schnittstellen zu externen Datenquellen bereit (Abb. 20h). Wie in 3.4.2 beschrieben, können dies sowohl Datenbanken wie z. B. Melderegister, KFZ-Register, Gesichts- oder Biometriedatenbanken mit strukturierten Daten als auch Quellen wie Foren, Blogs, soziale Netzwerke und Microbloggingdienste sein, die noch auszuwerten sind. Ω kann diese Informationen bei der Analyse der Bild- und Sensordaten ausnutzen, um beispielsweise eine bessere Risikoabschätzung vorzunehmen oder den OperateurInnen Informationen über Personen einblenden zu können. Beispielsweise könnte die Zusammengehörigkeit von Personen, die aus den Beziehungen in einem so-

zialen Netzwerks abgeleitet wurden, in der Darstellung eingeblendet werden. Neben Schnittstellen zur Erlangung von Daten, stellt Ω außerdem Schnittstellen für den Zugriff auf die eigenen Daten nach außen bereit (Abb. 20l). Diese sind vorgesehen für die Skalierbarkeit und Vernetzung einzelner Maßnahmen sowie die externe Auswertung der Daten, etwa zur Beweissicherung oder Fahndung.

4.3 Datenverarbeitung

Kameras übernehmen Bildverarbeitung

Bei Ω kommen zur dezentralen Datenverarbeitung Kameras mit integrierter Rechenleistung zum Einsatz (Abb. 20b). Man ist bestrebt, möglichst viel der Bildanalyse auf der Kamera oder sehr nahe am Ort der Erzeugung der Bilder durchzuführen. Auf jeder Kamera läuft ein *Multi-Object Tracker*, der alle Objekte verfolgt, und so lange sie sich im Bild befinden, eine kamerainterne temporäre Identifikationsnummer (*local-ID*) zuweist.[216] Die ermittelten Trajektorien werden abschnittsweise in kurzen Intervallen in Kombination mit Kamera-ID, *local-ID*, dem Ort des Objektes und der Zeit an einen zentralen *Multi-Camera Tracker* übertragen. Zusätzlich zu Trajektorien des Aufenthaltsortes von Personen sollen auch differenzierte Bewegungen des ganzen Körpers und des Gesichts erfasst und anhand dieser Daten Aktivität, Interaktion, Gang, Körpersprache (z. B. Aggressivitätslevel), Mimik und Gestik analysiert werden. Verknüpft mit der *local-ID* werden außerdem vermutliches Geschlecht und Alter sowie mit einer Person assoziierte Gegenstände, Gruppen oder Personen übertragen. Außerdem sind die Kameras mit Mikrofonen verknüpft, deren Tonspuren nach Mustern durchsucht werden können. Un-

216 Vgl. **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces", S. 2.

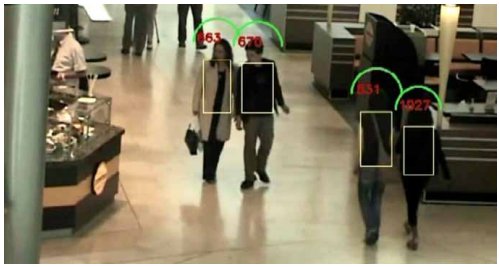


Abb. 21: Object Tracker versehen Objekte und Personen mit IDs.

ter Verwendung mehrerer verteilter Mikrofone kann über Triangulatur die Quelle von Geräuschen verortet werden.

Multi-Camera Tracker

Ein *Multi-Camera Tracker* setzt die Daten und Trajektorienstücke der einzelnen Kameras in ein globales Koordinatensystem ein, setzt sie anhand ihrer Geometrie zusammen und ordnet die temporären lokalen IDs globalen IDs zu.[217] Auf diese Weise werden die Bewegungen einzelner Personen über mehrere Kameras und ein ausgedehntes Gebiet verfolgt. Die Daten werden dann in einer *Trajektoriendatenbank* abgelegt, die Teil einer Datenbank für weitere Daten ist (Abb. 20k).

Verfolgung anhand von Personenmerkmalen

Ω besitzt die Fähigkeit, visuelle Personenmerkmale zu erfassen und für die Verfolgung über mehrere Kameras zu nutzen. Die Erfassung von *features* wie markante Farbkombinationen der Kleidung oder Körpermaße erlaubt eine wesentlich robustere kameraübergreifende Verfolgung. Bei lückenhafter Raumabdeckung wird großflächige Verfolgung durch derartige Informationen überhaupt erst ermöglicht, da die Personen in ver-

217 Vgl. ebd., S. 2.

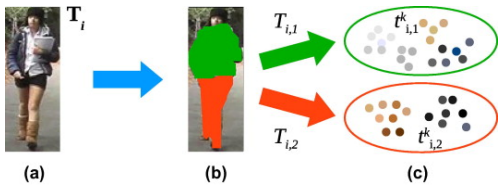


Abb. 22: *Appearance-based re-identification* unterteilt den Körper und speichert einen Vektor sich unterscheidender Farben. Diese werden mit einem Datensatz weiterer Vektoren zur Wiedererkennung abgeglichen.

schiedenen Szenen in Echtzeit wiedererkannt werden sollen (Abb. 22).[218] Um Personen auch über längere Zeiträume verfolgen zu können, reichen derartige *schwache Identifikationsmerkmale* nicht aus. Ω wird daher zusätzlich in der Lage sein, Biometriedaten zu erheben, als personenspezifischen Datensatz zu speichern und mit externen oder eigenen Datenbanken abzugleichen. Algorithmen zur Identifizierung aus der Entfernung anhand des Ganges und zur Gesichtserkennung ergänzen sich und sollen in Kombination verschiedene Probleme der Bildauflösung, Beleuchtung und Verdeckung minimieren. Bereits vorhandene Biometriedatensätze von Personen können immer wieder um fehlende Informationen ergänzt werden und daraufhin zu besseren Erkennungsraten führen. Beispielsweise kann das Modell des sehr variablen Laufstils um neue Bewegungsdaten ergänzt und die Körperabmessungen der Gliedmaßen verfeinert werden, um die Erkennungsrate zu steigern.

Datenanalyse

Die ermittelten Bewegungsdaten werden in Echtzeit ausgewertet. Diese Aufgabe übernimmt ein zentrales *Analysemodul*, welches auf die *Trajek-*

218 **Satta**, **Fumera** ; **Roli**, „Fast person re-identification based on dissimilarity representations“.

toriendatenbank zugreifen kann.[219] Das Analysemodul kann gleichzeitig eine Vielzahl von Aufgaben und Analyseinstanzen verschiedener Art ausführen. Diese Aufgaben werden von *Software-Agenten* abgebildet, die Verhalten von Personen bzw. deren Trajektorien nach übergeordneten Kriterien wie Geschwindigkeit, Richtung, Richtungsänderung, Umblicken oder abruptem Stehenbleiben untersuchen. Zu diesem Zweck können Anfragen nach geometrischen Strukturen an die *Trajektoriendatenbank* gestellt werden. Neben Einzelpersonen sollen ebenfalls Personengruppen erkannt und deren Bewegungen ausgewertet werden. Auch relative Trajektorien zwischen Objekten und Personen können analysiert werden, um Konzepte wie Meidungsverhalten, Folgen, Fliehen oder Anführen zu entdecken.[220] Anhand von Konvergenz- und Divergenzpunkten sollen Anhaltspunkte für eventuelle Konflikte (Personen laufen aufeinander zu), oder auch mögliche Gefahrenquellen (vor denen mehrere Personen davonlaufen) automatisch erkannt werden.

Eine weitere Aufgabe der *Agenten* kann sein, die beweglichen Kameras so zu steuern, dass auch bei unvollständiger Abdeckung der Bereiche möglichst vollständige Trajektorien erzeugt werden können, oder, dass optimales Bildmaterial vorliegt, um automatisch ein 3D-Modell einer verdächtigen Person zu generieren. Um Konflikte zu vermeiden, können die Agenten unterschiedliche Prioritäten haben. Wird beispielsweise eine Person, weil sie verdächtiges Verhalten gezeigt hat, von einem Agenten mit hoher Priorität verfolgt, so werden die Kameras bevorzugt auf diese Person ausgerichtet.[221] Ein solcher Agent kann bei Bedarf auch das Verfolgen der Person durch eine Drohne einleiten.

---

219 Vgl. **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces", S. 5.

220 Vgl. **Sester** ; **Kuntzsch**, *Szenenanalyse – Mustererkennung in Personen-Tracks*.

221 Vgl. **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces", S. 5.

Dezentrale Analyse

Um dem Ziel einer vollständigen Dezentralisierung nachzukommen, wird Ω außerdem „Taskmigration im Netz“ ermöglichen.[222] Das dezentrale Analysemodul soll damit weitestgehend entlastet werden, indem die Agenten auf den Prozessoren der Kameras laufen, und je nach Kapazität und Datenbedarf, von Prozessor zu Prozessor bzw. von Kamera zu Kamera migrieren können. Werden auf den Kameras selbst von normalem bzw. erwünschtem Verhalten abweichendes, oder konkret unerwünschtes Verhalten erkannt, können die Kameras in Ω selbstständig Alarme generieren, mit assoziierten Informationen anreichern und an geeignete Stelle weiterleiten. Empfänger der Alarme muss dabei nicht unbedingt Sicherheitspersonal sein. Alarme können auch anderen Agenten als Impuls zur weiteren Analyse dienen. Mit Techniken anderer Bereiche der Informatik z. B. *Peer-to-Peer-Kommunikation* oder verteilten Datenbanken könnten perspektivisch zentrale Komponenten und Datenvorräte abgeschafft werden. Ω kann daher in hohem Maße in der Anzahl der Kameras und der überwachten Fläche skalieren und zeitlich dynamisch erweitert und mit anderen kompatiblen Systemen gekoppelt werden.

4.4 Verhaltensmodelle

Ω soll mit einem minimalem manuellen Konfigurationsaufwand die Aufgabe zuverlässig ausführen.[223] Damit Ω möglichst generisch ist, werden zur Anpassung an einen konkreten Kontext spezialisierte Algorithmen und Module vorgehalten, die nach dem Plug-in-Prinzip integriert werden können.[224] Auch diese Module machen noch keine genauen Vorgaben

222 Vgl. **Hähner**, **Grenz** ; **Jänen**, *Verteilte vernetzte Kamerasysteme zur in-situ Erkennung Personen-induzierter Gefahrensituationen*.

223 Vgl. **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces“, S. 2.

224 Vgl. **Ko**, „A Survey on Behavior Analysis in Video Surveillance Applications“, S. 280.

von Modellen sicherheitskritischen oder unerwünschten Verhaltens.[225] Diese sollen im operativen Einsatz von Domänenexperten (z. B. dem Sicherheitspersonal) definiert oder automatisch erstellt werden. Dazu stehen neben Möglichkeiten manueller Erstellung von Regeln auch Algorithmen für unbeaufsichtigtes Lernen zur Verfügung. Auch während des Betriebes sollen stetig Daten wie Laufwege, Personenaufkommen sowie Informationen aus Sensoren gesammelt und Statistiken und Modelle angepasst werden.

4.5 Mensch-System-Interaktion

Die Darstellung der Kamerabilder und Analyseergebnisse für die OperateurInnen von Ω hat nicht mehr viel gemein mit einer Matrix von Kamerabildern wie in Abb. 2 dargestellt, sondern erinnert optisch und in der Steuerung eher an Computerspiele der Genres *Lebenssimulation* und *Echtzeitstrategiespiel*. Zentrales Element der Interaktion und der Informationsdarstellung ist eine interaktive virtuelle 3D-Umgebung, die Videobilder, 3D-Modelle und Informationen in einer einzigen umfangreichen Darstellung fusioniert (Abb. 20j). Das Betrachten von rohen Videobildern einzelner Kameras wird nur noch auf explizitem Wunsch ermöglicht, denn alle aufgenommenen Bildströme werden in Echtzeit perspektivisch in die 3D-Umgebung eingebettet. In dieser Übersichtsdarstellung werden als relevant eingestufte Analyseergebnisse, Statistikdaten, Laufwege etc. automatisch eingeblendet und Personen mit abweichendem Verhalten vom System markiert. Alarme werden am Ort ihres Aufkommens oder verknüpft mit Personen abgebildet. Vom System ermittelte Informationen über Personen wie Identität, Alter, Geschlecht, Laufwege, Agressivitätslevel und Assoziationen mit anderen Personen, Gruppen und Gegenstän-

225 Vgl. **d'Angelo** et al., „CamInSens : An Intelligent in-situ Security System for Public Spaces", S. 4.

den können textuell oder grafisch dargestellt werden. Auch akustische Ereignisse können in der 3D-Umgebung am entsprechenden Ort ihres Geschehens visualisiert werden. Da zu viele Informationen erhoben und durch Analyse generiert werden, als dass die OperateurInnen diese erfassen könnten, schätzt Ω selbstständig ein, welche der Informationen relevant sind. Ω hebt also genau die Informationen bzw. Personen hervor, die als eine Gefahr für den im Raum angestrebten Zustand klassifiziert wurden und gibt explizite oder über die Darstellungsweise implizierte Hinweise, wie die Situation zu interpretieren ist und wie mit ihr umgegangen werden sollte. Ω bietet dabei nicht nur die Darstellung und Annotation der aktuellen Geschehnisse an, sondern kann auch Bilder und Informationen aus der unmittelbaren oder weiter zurückliegenden Vergangenheit liefern. Wurde nach der Analyse der Daten ein automatisch generierter Alarm ausgelöst, kann Ω Szenen oder Informationen präsentieren, die nach der Interpretation des Systems zu dem Alarm geführt haben.

5 Gesellschaftliche Probleme und Auswirkungen

Nachdem die grundsätzlichen Funktionsweisen, das Zusammenspiel der einzelnen Komponenten und denkbare Fähigkeiten beschrieben wurden, sollen nun Probleme automatisierter Videoüberwachung identifiziert werden, die sich auf betroffene Individuen und Gesellschaft auswirken. Dabei wird sich immer wieder auf das System Ω bezogen. Es werden außerdem Ansätze zur Reduzierung oder Vermeidung der Probleme diskutiert.
Die Probleme sind zum Teil stark miteinander verwoben und bedingen sich gegenseitig. Die Kapitel wurden in eine möglichst logische Reihenfolge gebracht. Es kommt jedoch vor, dass einzelne Aspekte eines Problems erst in einem späteren Kapitel hergeleitet und erwähnt werden. Diese Aspekte sind jedoch für die Argumentation innerhalb eines Kapitels nicht erforderlich, sondern ergänzen diese lediglich.

5.1 Datenschutz und *privacy*

In Kapitel 2.3.2 wurde die Bedeutung von Datenschutz und *privacy* bei *manueller Videoüberwachung* dargelegt.
Mit dem Einsatz von Systemen wie Ω wird die Erhebung von Daten sowie die Quantität und Qualität der automatisierten Verarbeitung erheblich ausgeweitet. Mit diesen erweiterten Möglichkeiten der Verarbeitung ändert sich auch der Charakter der Videoüberwachung.
Da die genaue Ausgestaltung von Systemen wie Ω nicht genau vorhersehbar ist, können nur ausgewählte Aspekte und grundsätzliche Implikationen der Automatisierung für den Datenschutz und das Recht auf informationelle Selbstbestimmung vorausgedacht werden.

Von Raumüberwachung zu Überwachung von Individuen

Während *manuelle Videoüberwachung* durch einzelne Maßnahmen nur Raumüberwachung und keine Überwachung von Einzelpersonen darstellte[226], stellt Ω ein mögliches Werkzeug zur Sammlung von Daten von Individuen, also zur Überwachung von Einzelpersonen dar. Ω kann Personen anhand von Gang, Gesicht und Kleidung wiedererkennen, Laufwege über mehrere Kameras und sogar potentiell systemübergreifend verfolgen, außerdem weitere Daten erheben und mit bestimmten Personen verknüpft für unbestimmte Zeit speichern. Über zeitlich rückwärts gerichtete Überwachung durch Auswertung von Rohdaten kann eine Identifizierung und Informationsgewinnung jederzeit nachträglich vorgenommen werden. Daran, dass dies nur eine Frage der Rechenkapazität ist, wird noch einmal deutlich, warum nicht zwischen Personenbezogenheit und Personenbeziehbarkeit unterschieden wird.

Verkettung und Zweckgebundenheit der Daten

Nach § 6b des BDSG müssen für jede einzelne Maßnahme konkrete Zwecke festgelegt werden. Durch die Vernetzung der Kameras, die Verknüpfung der Kameranetze oder die Abfrage der Datenbestände von Außen verschwimmen die Grenzen zwischen den einzelnen Überwachungsmaßnahmen. Übermittlung und dezentrale Verarbeitung erhöht die Wahrscheinlichkeit, dass die Zweckgebundenheit verletzt wird. Selbst wenn jede einzelne Maßnahme begründet und legal ist, muss es die Verbindung dieser Netze und die Verarbeitung der Daten in einem anderen Rahmen nicht unbedingt sein.[227] Es besteht außerdem eine erhöhte Wahrscheinlichkeit, dass durch die Verkettung von einzelnen, vermeintlich „belanglosen" Daten diese einen erhöhten Stellenwert bekommen. Auch in die-

226 **Klauser**, *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*, S. 91.

227 **Coudert** ; **Dumortier**, „Intelligent Video Surveillance Networks: Data Protection Challenges", S. 979.

sem Punkt spiegelt sich die Feststellung des Volkszählungsurteils wider, dass unter den Bedingungen der automatischen Datenverarbeitung kein Datum mehr als „belanglos“ angesehen werden kann. Allein anhand von Aufenthaltsorten einer Person, können umfangreiche Rückschlüsse über sie und ihre Aktivitäten gezogen werden. Dies wurde anhand der Vorratsdaten des Grünenpolitikers Malte Spitz, die er von der Telekom eingeklagt hatte, auf Zeit.de eindrucksvoll veranschaulicht.[228]

Einsichtnahme und Überprüfbarkeit durch Betroffene

Nach der *Charta der Grundrechte der Europäischen Union* Artikel 8 (2) und auch dem Bundesdatenschutzgesetz § 19 hat jede Person das Recht, Auskunft über die sie betreffenden erhobenen Daten zu erhalten und die Berichtigung der Daten zu erwirken. Sind die personenbezogenen Daten weder automatisiert noch in nicht automatisierten Dateien gespeichert, wird nach dem BDSG die Auskunft nur erteilt, „soweit der Betroffene Angaben macht, die das Auffinden der Daten ermöglichen, und der für die Erteilung der Auskunft erforderliche Aufwand nicht außer Verhältnis zu dem vom Betroffenen geltend gemachten Informationsinteresse steht“. In Ω liegen die personenbeziehbaren Daten als Rohdaten so vor, dass eine Identifizierung jederzeit durchgeführt werden könnte. Dass Betreiber jedoch Schnittstellen für Anfragen der Betroffenen bereithalten, die den Aufwand genügend reduzieren, ist unwahrscheinlich. Den Betroffenen fehlt damit jede reale Möglichkeit, Einsicht in die über sie gespeicherten Daten zu nehmen oder eine Korrektur oder Löschung zu erzwingen. Auch eine mögliche Weitergabe, Speicherung, Verarbeitung oder Nutzung kann von Betroffenen nicht überprüft werden. In Deutschland ist in vielen Fällen aktuell stattfindender Überwachungsmaßnahmen

228 **Zeit Online (Hrsg.):** *Verräterisches Handy*.

gar nicht ersichtlich, wer der Betreiber ist, da die Kennzeichnungspflicht nach BDSG § 6b (2) oft nicht eingehalten wird.[229]

Datensicherheit

Auch die Datensicherheit der personenbezogenen Daten kann von Betroffenen nicht überprüft werden. Betroffene müssen daher befürchten bzw. können nicht ausschließen, dass Daten durch „leaking"[230] von innerhalb des Systems nach außen gebracht werden, oder durch „tapping"[231] von außerhalb des Systems extrahiert werden. Auf Datensicherheit von einem so komplexen System wie Ω, kann an dieser Stelle nicht weiter eingegangen werden.

Diese Umstände einer mangelnden Kontrolle bei gesteigerten Verarbeitungsmöglichkeiten stellen gegenüber *manueller Videoüberwachung* einen viel tieferen Eingriff in das Recht auf informationelle Selbstbestimmung dar und führt zu einem starken Informationsungleichgewicht zwischen Überwachenden und Betroffenen. Die Effekte, die dieses Ungleichgewicht auf Betroffene hat, werden in Kapitel 5.4 behandelt.

Das Ungleichgewicht und der Rechtseingriff wird weiter verstärkt durch den inhärent hohen Datenbedarf von automatisierter Videoüberwachung, der in Spannung mit dem Datenschutzprinzip der Datensparsamkeit steht. Dieser Datenbedarf wird im Folgenden anhand technischer und konzeptueller Gegebenheiten hergeleitet.

5.1.1 Spannung zwischen Datenbedarf und Datensparsamkeit

Die Prinzipien der Datenvermeidung und der Datensparsamkeit nach § 3a BDSG gelten auch für das Durchführen von Videoüberwachung.[232]

229 **Wahlbrink**, *Zahlreiche Rechtsverstöße bei der Videoüberwachung : Wahlbrink: Behörden und Kommunen ignorieren Datenschutzgesetz*.

230 **Moncrieff**, **Venkatesh** ; **West**, „Dynamic Privacy in Public Surveillance", S. 25.

231 **Senior**, *Protecting Privacy in Video Surveillance*, S. 92.

232 **Weichert**, „Private Videoüberwachung und Datenschutzrecht".

Demnach ist „die Erhebung, Verarbeitung und Nutzung personenbezogener Daten und die Auswahl und Gestaltung von Datenverarbeitungssystemen [...] an dem Ziel auszurichten, so wenig personenbezogene Daten wie möglich zu erheben, zu verarbeiten oder zu nutzen."
Im Folgenden wird hergeleitet, warum bei Systemen wie Ω die Funktionsweise der Technik und das Ziel der automatisierten Überwachung mit den Prinzipien der Datensparsamkeit und der Datenvermeidung in Spannung und sogar im Widerspruch steht. Dazu wird das grundlegende Konzept automatisierter Videoüberwachung analysiert.

Vorhersehbarkeit

Die Bedeutung, die den Techniken für Sicherheit und Überwachung heute beigemessen wird, ist unter anderem auf den Glauben zurückzuführen, dass Überwachung der Zukunft von Menschen möglich ist.[233] Der hier verfolgte präventive Ansatz beinhaltet die Grundannahme, dass im visuell beobachtbaren Verhalten von Personen Informationen vorliegen, anhand derer das Aufkommen von unerwünschten Geschehnissen im Allgemeinen voraussagbar und im Speziellen algorithmisch berechenbar sind.
Mit dem *Blacklist-* und dem *Whitelistansatz* wird versucht, von konkretem oder abweichendem beobachtbaren Verhalten auf unerwünschte Geschehnisse zu schließen. Dieser Versuch erscheint anhand folgender drei Tatsachen jedoch wenig erfolgversprechend: *Erstens* muss Auffälligkeit nicht auf Gefahr oder Kriminalität hindeuten; *zweitens* muss Gefahr und Kriminalität nicht beobachtbar von „normalem" Verhalten und Geschehnissen abweichen und *drittens* sind Verhalten und Geschehnisse nicht zwingend eindeutig interpretierbar. Menschen besitzen mitunter eine beachtenswerte Fähigkeit, ihr Verhalten so anzupassen, dass ihr Anliegen nicht detektiert werden kann.[234]

233 **Amicelle**, „Exclusion and discrimination", S. 220.
234 **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 19.

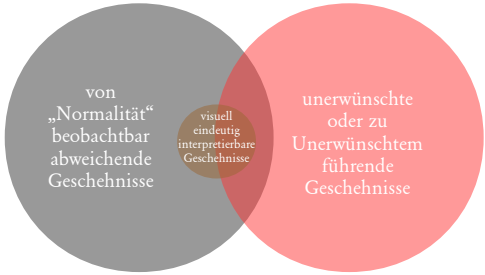


Abb. 23: Venn-Diagramm zur Veranschaulichung, dass Abweichung von Normalität nicht Unerwünschtheit impliziert, Unerwünschtheit nicht Beobachtbarkeit impliziert und Beobachtbarkeit nicht Eindeutigkeit impliziert.

Diese drei Implikationen werden in Abb. 23 mit einem Venn-Diagramm über Mengen von Geschehnissen veranschaulicht.
Daraus ergibt sich, dass *Blacklist-* und *Whitelistansatz* bereits in der Theorie weder vollständig noch korrekt Gefahr und Kriminalität detektieren oder voraussagen können. Setzt man die Verfahren dann praktisch ein, verschlechtert sich Vollständigkeit und Korrektheit noch mehr, da ungenaue Ausgangswerte und komplexitätsreduzierende Mechanismen genutzt werden. In Kapitel 2.3.1 wurde beschrieben, dass bei Videoüberwachung ein Transfer der Kontrolle vom eigentlichen auf einen abstrakten Raum stattfindet. Zusätzlich zu den oben beschriebenen technischen Filtern (etwa durch Wahl der Kameraposition) findet mit Ω eine weitere Informationsreduktion während der Bildverarbeitung und der Analyse mit künstlicher Intelligenz statt. Deren grundsätzlichen Konzepte – Modellierung und Wahrscheinlichkeitsrechnungen – führen also zu weiterem Informations- und Komplexitätsverlust und zu Informationsverzerrung.

Das Ziel, trotz dieser Limitierung eine möglichst hohe Effektivität und Effizienz zu erreichen, impliziert eine bestimmte Gestaltung und Konfiguration von Systemen wie Ω.
Im Folgenden soll die Entscheidungsfindung von Ω betrachtet werden. Es können zwei entscheidende Aspekte festgestellt werden. *Erstens* muss Ω nach dem Kosten-Nutzen-Prinzip zum Nachteil der Betroffenen so konfiguriert werden, vorwiegend zur Verdächtigung, zu eingehenderer Beobachtung zu tendieren, wodurch mehr Daten erhoben werden. *Zweitens* hat Ω um Fehlalarme zu vermeiden, einen enormen Informationsbedarf. Ω steht daher in Spannung mit dem Prinzip der *Datensparsamkeit* und hat somit einen inhärenten datenschutzfeindlichen Charakter.

Entscheidungsfindung

Während des Betriebes von Ω werden bei der Analyse und Bewertung der Geschehnisse auf verschiedensten Ebenen automatisch Entscheidungen getroffen, die (bei einer nicht voll automatisierten Videoüberwachung) in letzter Instanz zu einer Information für OperateurInnen wird. Dies können binäre Informationen sein wie „Alarm", oder aber differenziertere Informationen wie das Aggressivitätslevel einer Person als Prozentsatz oder nach dem Ampelprinzip diskretisiert.
Beispielhaft und vereinfachend soll im Folgenden die binäre Entscheidung zwischen „Alarm wird angezeigt" und „es wird kein Alarm angezeigt" betrachtet werden.
Damit Ω effektiv und effizient ist, müssen die richtigen Entscheidungen getroffen werden. Es sollen möglichst alle relevanten Geschehnisse rechtzeitig detektiert werden und möglichst keine Fehlalarme erzeugt werden. Ω soll also *vollständig* und *korrekt* arbeiten.

Fehlentscheidungen

Aus der fehlenden Eindeutigkeit der Geschehnisse, der Komplexitätsreduzierung und Informationsverzerrung kommt es jedoch zwangsläufig

zu *Fehlalarmen* und fälschlicherweise nicht ausgelösten Alarmen (*miss*). Diese beiden stehen im unmittelbarem Zusammenhang, der im Folgenden dargelegt wird.

Die Analyse von schwer entdeckbaren Signalen, hier die Erkennung unerwünschter Geschehnisse im beobachteten Raum, ist Gegenstand der von John Swets und David Green entwickelten Signalentdeckungstheorie (SDT).[235] Sie stellt ein Qualitätsmaß für Erkennungsleistung bereit. Es werden vier Fälle je nach Vorhandensein und Erkenntnis einer Gefahr unterschieden: Verpasser (*miss*), korrekte Ablehnung (*correct rejection*), Treffer (*hit*) und falscher Alarm (*false alarm*) (Tabelle 1).[236]

	eine Gefahr erkannt	**keine Gefahr erkannt**
Gefahr vorhanden	Treffer	Verpasser
keine Gefahr vorhanden	falscher Alarm	korrekte Ablehnung

Tabelle 1: Fallkombinationen nach der Signalentdeckungstheorie nach Swets und Green

Falschakzeptanz-, Falschrückweisungsrate und Effizienz

Für die verschiedenen Fälle können mit Testdatensätzen oder beim echten Einsatz relative Häufigkeiten ermittelt werden – die *Falschakzeptanzrate* (FAR) und die *Falschrückweisungsrate* (FRR) bzw. die mit der FRR zusammenhängende *Trefferrate*. Das Sensitivitätsmaß d' ist die Differenz der z-transformierten FAR und Trefferrate und gibt die Qualität der Erkennung bzw. deren *Effizienz* an:

$$d' = z(Trefferrate) - z(FAR).$$

Je geringer die FAR und größer die Trefferrate (bzw. kleiner die FRR) desto größer die Sensitivität und die Effizienz beim Einsatz.

235 **Green** ; **Swets**, *Signal Detection Theory and Psychophysics*.
236 Ebd., S. 34.

Eine weitere relevante Größe ist die Antworttendenz (auch Reaktionsneigung genannt) c:

$$c = -0.5 * (z(FAR) + z(Treffrate)).$$

Ohne die Korrektheit des Alarms zu berücksichtigen gibt c an, wie sehr dazu tendiert wird, einen Alarm zu geben. In der Literatur werden Systeme mit hoher Reaktionsneigung mitunter ebenfalls als *sensitiv* bezeichnet. Um die beiden Größen zu unterscheiden, soll hier stattdessen das Adjektiv *empfindlich* und das Substantiv *Empfindlichkeit* für hohe Antworttendenz und je nach Kontext *sensitiv* oder *effizient* für hohe Sensitivität genutzt werden.

Zusammenhang FAR und FRR

Vollständigkeit und *Korrektheit* der gegebenen Alarme sind sich gegenseitig komplementär beeinflussende Ziele. In Ω manifestiert sich die Regulierung dieses Zusammenhanges technisch auf verschiedensten Abstraktionsebenen im Konzept von *Schwellwerten* beim Fällen von (binären) Entscheidungen. Ein einziger Schwellwert beeinflusst sowohl die Falschrückweisungsrate als auch die Falschakzeptanzrate eines Algorithmus. Je geringer der Schwellwert, desto kleiner die FRR und gleichzeitig größer die FAR. Man erhält in diesem Fall eine große Menge Fehlalarme und wenige *miss*. Erhöht man den Schwellwert, so nimmt die Anzahl der *miss* zu, die der Fehlalarme nimmt ab.

Über den Schwellwert kann ein Arbeitspunkt eingestellt werden, der das Verhältnis von FAR und FRR festlegt. Dazu wird der Grad der Auswirkungen (Kosten) im *miss*-Fall und bei falschem Alarm ins Verhältnis gesetzt. Ob beim Einsatz von Ω also ein Fehlalarm im negativsten Fall zum Ignorieren der Warnung, einer Befragung einer verdächtigten Person oder zum Aktivieren einer Selbstschussanlage in einem Grenzgebiet

führt, hat entscheidende Auswirkungen auf die Wahl der in Ω vorkommenden Schwellwerte.

Tendenz zum Misstrauen und Kosten

Im Falle der Videoüberwachung unterscheiden sich die Kosten im Einzelfall für *miss* und *Fehlalarm* signifikant. Die immensen finanziellen Kosten und der Verlust von Menschenleben bei einem terroristischer Anschlag, der wegen zu geringer Empfindlichkeit nicht rechtzeitig erkannt wurde, werden im konkreten Fall den augenscheinlich geringeren Kosten präventiver Festnahmen verdächtiger Personen gegenübergestellt. Bei ungerichteter Überwachung zur Verhinderung von potentiellem Terrorismus müssten die Systeme besonders empfindlich eingestellt werden.[237] Es ist also bei der Konfiguration von Ω davon auszugehen, dass Arbeitspunkte in hohem Maße auf Kosten der FAR gewählt werden. Nach dem oben gewählten Beispiel von Alarmen würde dies zu einer großen Anzahl von Fehlalarmen führen. Allgemeiner betrachtet heißt dies jedoch, dass nicht nur Entscheidungen über Alarme sondern auch systeminterne Entscheidungen eher misstrauisch gefällt werden. Je nach Entscheidungsebene geschieht dies auf Kosten der Rechte der Betroffenen. Durch hohe Empfindlichkeit des Systems werden tendenziell mehr Daten erhoben, verarbeitet und durch detailliertere Analyse und Verkettung neue Daten erzeugt. Wird beispielsweise ein *geringer* Schwellwert für Aggressivität der Körpersprache gewählt, so wird ein Agent, der Kameras auf diese Person ausrichtet, aus geringerem Anlass instanziiert und werden somit häufiger zusätzliche Daten erhoben und verarbeitet als bei höherem Schwellwert.

Sensitivitätsgewinn durch Optimierung der Algorithmen

Ist die Sensitivität auch nach Optimierung des Arbeitspunktes noch nicht zufriedenstellend, ist ein mögliches Mittel die Verbesserung der Algorith-

237 **Čas**, „The relevance of social and economic costs of surveillancy – Conclusion", S. 253.

men. Die Sensitivität bei gleicher Datenbasis zu erhöhen, wird fast sportlich betrieben. Beispielsweise werden Algorithmen zur Gesichtserkennung in wissenschaftlichen Wettbewerben mit Musterdatenbanken miteinander verglichen und immer subtilere Optimierungsansätze verfolgt. Über maschinelles Lernen und Datenbanken mit Trainingsdaten wird versucht, die Algorithmen automatisiert zu optimieren. Diese Optimierung hat jedoch auf Grund der Datenbasis und der technischen Möglichkeiten seine Grenzen. Selbst mit bestmöglichem Algorithmus kann nur eine begrenzte Sensibilität erreicht werden.

Sensitivitätsgewinn durch zusätzliche Informationen

Die Sensitivität kann potentiell nur dann noch weiter gesteigert werden, wenn die Menge der berücksichtigten Informationen erhöht wird. Ein Beispiel ist die ungenügende Erkennungsrate von Gesichtserkennungsalgorithmen bei unvorteilhaften Lichtverhältnissen oder Kameraperspektiven. Zur Steigerung der Erkennungsrate wird die Technik um die Identifikation am Laufstil ergänzt und dadurch mehr Informationen erhoben und verarbeitet.

Die oben beschriebene theoretische Limitierung, die sich aus der Uneindeutigkeit der Interpretation von Verhalten und der möglichen Unbeobachtbarkeit von Unerwünschtem ergibt (Abb. 23), ist mit hoher Wahrscheinlichkeit dafür verantwortlich, dass die Effektivität von Systemen wie Ω auch nach Optimierung der Schwellwerte, der Algorithmen und der Datenauswahl für die Interessen der Betreiber nicht ausreichen. Es ist daher mit der Einführung von zwei weiteren Maßnahmen zu rechnen. Diese sind *risk profiling* und das Ausnutzen von *externen Informationen*, die über Informationen, die aus dem beobachtbaren Verhalten erkannt werden können hinaus gehen.

Risk profiling

Beim *risk profiling* wird im Gegensatz zu bisher beschriebenen Methoden ein Risiko nicht am tatsächlichen Verhalten oder individuellen Informationen über eine Person abgeschätzt, sondern anhand von hypothetischen Informationen, die von angesammelten Daten mit statistischen Methoden abgeleitet wurden. Während also das vermeintliche Aggressivitätslevel einer Person anhand individueller „Handlung" und Information ermittelt wird, ist die Abschätzung des Risikos aufgrund bestimmter Kleidungsstücke, der Hautfarbe oder der Herkunft *risk profiling*. Auf diese Weise wird versucht, das Gefahrenpotential von Menschen auch dann einzuschätzen, wenn diese kein auffälliges, nicht hinreichend auffälliges oder uneindeutiges Verhalten zeigen. Auf *risk profiling* wird im Zusammenhang mit Diskriminierung in Kapitel 5.3 näher eingegangen. An dieser Stelle sei hervorgehoben, dass auch für *risk profiling*, die Datenerhebung und -verarbeitung erweitert werden muss.

Externe Informationen

Da *risk profiling* auf Statistiken und Kategorisierung beruht, kann auch dieses zu Fehlalarmen bzw. Fehleinschätzungen führen und wahrscheinlich nicht die erhoffte Effizienz erbringen. Das Forschungsprojekt *INDECT* ergänzt daher die Videoüberwachung und Risikoabschätzung zusätzlich mit Informationen aus anderen Quellen, wie in 3.4.2 dargestellt wurde. Wie bereits erwähnt, kann hierauf nicht weiter eingegangen werden. Betont sei in diesem Zusammenhang, dass durch die automatisierte Verkettung und Verarbeitung von Informationen von verschiedensten Stellen der Rechtseingriff entschieden vertieft würde.

Zusammenfassung

Zusammenfassend kann festgestellt werden, dass das Interesse und die Funktionsweise automatisierter Videoüberwachung dem Prinzip der Datensparsamkeit konträr gegenübersteht. Aufgabe der Überwachung des

öffentlichen Raums ist es, Informationen aus einer komplexen Welt in Kategorien wie *erwünscht* und *unerwünscht* einzuteilen oder zur nachträglichen Aufklärung von Geschehnissen vorrätig zu halten. Welche Informationen dazu von Nutzen sein könnten, ist vorher nicht klar. Für rückwärts- und vorwärtsgerichtete Überwachung werden so viele Informationen gesammelt wie möglich. Die rückwirkende Analyse von Laufwegen im Nachhinein verdächtig wirkender Personen beispielsweise, benötigt Daten, die Minuten, Stunden, Tage oder sogar Monate in der Vergangenheit liegen. Die Steigerung der Effizienz einer Risikoabschätzung erfolgt nicht nur auf Kosten der Datensparsamkeit, sondern steigert den Datenbedarf enorm. Außerdem ist damit zu rechnen, dass Systeme wie Ω tendenziell misstrauisch konfiguriert werden und damit Personen eher eingehender überprüft werden und daher mehr Daten erhoben und ausgewertet werden.

Dass verdachtsunabhängige, ungerichtete Datenspeicherung auf Vorrat auch gängige Überwachungspraxis von demokratischen Ländern ist, wurde in den letzten zwei Jahren am Internetüberwachungsprogramm *Tempora* des britischen Geheimdienstes und *PRISM* der NSA deutlich. Der Internetverkehr wird abgefangen und vorrätig zur Analyse gespeichert. Nach Medienberichten geschehe dies bei *PRISM* für drei bis sechs Monate – Metadaten würden für immer gespeichert werden.[238] Nicht nur Geheimdienste verfolgen dieses Prinzip. Auch in der EU sollen nach der Richtlinie 2006/24/EG verdachtsunabhängig alle Kommunikationsmetadaten der BürgerInnen vorrätig sechs Monate lang durch die Internetprovider gespeichert werden. Die deutsche Umsetzung der Richtlinie wurde vom Bundesverfassungsgericht 2010 in einem Urteil als verfassungswidrig erklärt wird jedoch immer wieder von PolitikerInnen eingefordert.[239]

238 **Wilkens**, *PRISM-Überwachung : BND und NSA in einem Boot*.
239 BVerfG, Urteil vom 2. März 2010, 1 BvR 256/08.

Der technisch und konzeptuell bedingte inhärente Datenbedarf für vorwärtsgerichtete Überwachung und Echtzeitauswertung, sowie das Datensammeln für rückwärts gerichtete automatisierte Überwachung verletzt die Datenschutzgrundprinzipien der Datensparsamkeit und der Datenvermeidung. Der Grundrechtseingriff und das Informationsungleichgewicht zwischen Betroffenen und Überwachenden wird dadurch verstärkt.

5.1.2 Diskussion der Datenschutzmaßnahmen

The system will probably grow, but there will be more and better protection of personal data.
PROJEKT-KOORDINATOR DES INDECT-FORSCHUNGSPROJEKTES

Im Vorangegangenen wurde hergeleitet, dass Systeme wie Ω einen massiven inhärenten Informationsbedarf haben, der in Spannung zu grundsätzlichen Prinzipien des Datenschutzes steht. Mit den in 3.6 beschriebenen Techniken sollen Datenschutzvorkehrungen getroffen werden. Hauptaugenmerk der Techniken wie *datahiding* oder Konzepten wie dem *Drei-Stufen-Modell* liegt darauf, die Daten der Betroffenen zu anonymisieren, um den Grundrechtseingriff möglichst gering zu halten. Nach §3 (6) des Bundesdatenschutzgesetzes ist Anonymisieren „das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können."

Datahiding

Der Ansatz des *datahiding* kann der Definition von Anonymisierung nicht gerecht werden. Zwar existieren Lösungen, die eine Unkenntlichmachung als sensibel erkannter Bereiche (ROI) im Bildmaterial zufriedenstellend durchführen, doch sowohl die Auswahl als auch die Verlässlichkeit der Erkennung der zu versteckenden Informationen müssen im

Sinne des Datenschutzes als ungenügend bewertet werden. Dies liegt an drei Aspekten: Zum Einen kann keine verlässliche Detektion der zu versteckenden Informationen gewährleistet werden, zum Anderen wird sich beim Ausblenden auf Biometrie und schwache Identifikatoren wie Körpermaße beschränkt[240], und außerdem ist damit zu rechnen, dass *datahiding* nur zum Schutz vor den Blicken der OperateurInnen eingesetzt wird und Daten trotz vermeintlichem Schutz personenbeziehbar vorliegen.

Verlässlichkeit der Detektion

Bei den meisten Ansätzen sind die Informationen (ROI), auf die die Maßnahmen abzielen, Körper oder Gesicht. Die Algorithmen können jedoch nicht garantieren, dass alle Menschen und Gesichter vollständig ermittelt werden. Der Grund hierfür liegt nicht an der Qualität der Algorithmen, sondern in der Natur der Sache. Verdeckungen, schlechte Lichtverhältnisse, aber vor allem Abweichung von Annahmen, wie ein Gesicht oder ein Mensch auszusehen haben, sind verantwortlich für die Aussetzer. In Zukunft könnten lichtempfindlichere Kameras genutzt und Überlappungsprobleme mit Bildern aus mehren Perspektiven gelöst werden. Eine zuverlässige Erkennung der ROI kann dann jedoch noch immer nicht garantiert werden. Das Problem des Abweichens vom als normal angenommenen Erscheinungsbild eines Menschen oder eines Gesichtes bleibt bestehen. Problematisch an der mangelnden Verlässlichkeit ist, dass schon ein einziges Frame, bei dem eine ROI ungeschützt vorliegt, die Identität oder andere zu schützende Informationen offenlegen könnte.[241] Alle mit der bisher anonymen Person verknüpften Daten sind nur durch dieses eine Bild mit einem Mal zu personenbeziehbaren Daten geworden. Dass

240 **Senior**, *Protecting Privacy in Video Surveillance*, S. 37.
241 Ebd., S. 38.

die Verfahren robust genug für Datenschutzanforderungen werden, ist für die nächsten Jahre nicht zu erwarten.[242]

Nichtsdestotrotz wird der Ansatz in wissenschaftlicher Literatur als tauglich eingestuft.[243] Es wird in den Vordergrund gerückt, dass die automatisierte Extraktion von privacy-relevanten Informationen und der Missbrauch durch OperateurInnen mit den vorhandenen Techniken erschwert wird.

Identifizierbarkeit anhand anderer Informationen

Dass das *datahiding* von Augenpartie oder Kopf oder Unschärfen des Körpers keinen sicheren Schutz vor Identifizierung leistet, ist offensichtlich (Abb. 18), da weitere Merkmale zur Identifizierbarkeit beitragen. Anhand von Identifizierung am Gang oder Kleidung können derart „geschützte" Daten trotzdem automatisiert durchsucht werden. Folglich werden, wie in Abb. 17 dargestellt, auch mehr als nur Augen und Kopf ausgeblendet oder unkenntlich gemacht. Neben diesen biometrischen Merkmalen bewirken jedoch auch andere Merkmale, dass Personen identifiziert werden können und Daten als *personenbeziehbar* gelten müssen.[244] Diese Merkmale können unterschiedlichster Natur sein: Gegenstände, Farbkombination der Kleidung oder aber im weitesten Sinne räumlicher und zeitlicher Kontext. Der Kontext kann selbst dann zu Identifizierbarkeit führen, wenn die Person im Bild nur durch einen Punkt ersetzt wird (Abb. 17). Die Tatsache, dass eine Person morgens zu einer bestimmten Zeit auf einem Bahnhof den Zeitungskiosk aufschließt, deckt ihre Identität auf. Über rückwärts gerichtete Überwachung können z. B. über Laufwege weitere Informationen mit der beobachteten Person verknüpft

242 **Senior**, *Protecting Privacy in Video Surveillance*, S. 30; **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?", S. 336.

243 **Senior**, *Protecting Privacy in Video Surveillance*, S. 39.

244 **Coudert** ; **Dumortier**, „Intelligent Video Surveillance Networks: Data Protection Challenges", S. 976.

werden. Liegen neben Videodaten auch andere Daten z. B. Prozesse an Geldautomaten oder Daten von elektronischem, personengebundenen Bezahlsystem für öffentlichen Nahverkehr vor, so müssen die Daten auch hierdurch als personenbeziehbar behandelt werden. Es ist nicht anzunehmen, dass Informationen wie in den beiden Beispielen nur von Menschen auswertbar sind. Sie können gezielt oder automatisiert gelernt, auch algorithmisch ausgewertet werden. Während Merkmale wie Gegenstände und Farben zusätzlich ausgeblendet werden können, ist die zeitliche und räumliche Komponente bei Videoüberwachung in Echtzeit nicht zu verbergen, ohne dass die Ausführbarkeit der eigentlichen Überwachungsaufgabe darunter leidet.

Selbst wenn also beim *datahiding* das Problem einer unvollständigen Detektion der ROI nicht ins Gewicht fallen würde, bestehen begründete Annahmen, dass *datahiding* kein ausreichendes Mittel zum Datenschutz darstellen kann. Das Finden effektiver Mittel gegen die Identifizierung auf alternativen Wegen erscheint unwahrscheinlich oder ist zumindest in absehbarer Zeit nicht zu erwarten.

Reversibles *datahiding*

Datahiding steht im Konflikt mit der eigentlichen Überwachungsaufgabe, möglichst viele Informationen auswerten zu können. Videoüberwachung soll in den meisten Fällen auch bei der Aufklärung von Straftaten helfen. Für eine mögliche Ermittlung müssen die Daten daher in irgendeiner Form zur Verfügung stehen. Für dieses Anliegen wurde reversibles *datahiding* entwickelt. Wenn die Daten jedoch nach wie vor personenbeziehbar vorliegen, kann nach dem Verständnis des BDSG nicht die Rede von Anonymisierung sein, denn von einem „unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft“ kann bei dem Akt der Entschlüsselung bzw. dem Eingeben eines Passwortes nicht ausgegangen wer-

den. Es handelt sich in diesem Fall um eine Maßnahme zur Datensicherheit, nicht zur Anonymisierung und zum Datenschutz.

Datenschutz wem gegenüber?

Während Kritik an Videoüberwachung bezüglich Datenschutz zunächst prinzipiell an der Erhebung der Daten geäußert wurde, scheint die öffentlichen Diskussion sich immer mehr dahin zu entwickeln, allein die OperateurInnen für den Verlust von *privacy* verantwortlich zu machen. Die Themen Voyeurismus, Diskriminierung und Leaks von Videomaterial durch Mitarbeiter, sind nicht nur Gegenstand von Diskussionen sondern auch der Forschung an vermeintlich datenschutzfreundlicher Technik geworden.[245] In *INDECT* werden Wasserzeichentechniken zur Rückverfolgung (nicht zur Verhinderung) von Leaks entwickelt; *datahiding* soll wie eben beschrieben die Identität vor OperateurInnen geheim halten; Beschränkung des Zooms und weitgehende Automatisierung sollen OperateurInnen von voyeuristischem Verhalten abhalten.

Alle diese Techniken und Möglichkeiten werde als Verbesserung des Datenschutzes gegenüber manuellen Systemen angepriesen.[246] Dass die Menge erhobener Daten, die der *manuellen Videoüberwachung*, wie in 5.1.1 hergeleitet, um ein Vielfaches übersteigt und diese Daten automatisiert verarbeitet, gespeichert und durchsucht werden können, bleibt bei dieser Überlegung weitestgehend unberücksichtigt. Die signifikant höhere Effizienz bei der Auswertung bedeutet jedoch einen tieferen Eingriff in die schutzwürdigen Interessen der Betroffenen.[247]

245 Vgl. **Cuxhavener Nachrichten (Hrsg.):** *Die „kleinen Brüder" schauen nie weg: Kamera-Überwachung in Cuxhaven*.

246 **Senior**, *Protecting Privacy in Video Surveillance*, S. 35 ff.

247 **Roßnagel**, **Desoi** ; **Hornung**, „Gestufte Kontrolle bei Videoüberwachungsanlagen : Ein Drei-Stufen-Modell als Vorschlag zur grundrechtsschonenden Gestaltung", S. 698.

Scheinbare Legitimität

Wird der Eingriff in das Recht auf informationelle Selbstbestimmung durch scheinbar datenschutzfördernde Techniken verschleiert, so wird dies bei der Abwägung der Verhältnismäßigkeit bei der Entscheidung über die Errichtung einer konkreten Maßnahme die Balance zwischen Eingriffstiefe, Tauglichkeit und Notwendigkeit stark verfälschen. Wenn die Persönlichkeitsrechte scheinbar geschützt werden, besteht die Gefahr, dass die Installation von Videoüberwachungsanlagen auch an Orten legitim scheint, an denen sie unvertretbar ist.

5.2 OperateurInnen als Teil des Automatismus

Mit Automatisierung von Videoüberwachung wird angestrebt, einzelne Aufgaben der OperateurInnen zu automatisieren oder die OperateurInnen ganz zu ersetzen. Ω übernimmt nicht nur die Auswahl der Bilder, sondern auch deren Interpretation und gibt explizite und implizite Hinweise, wie mit einer vermeintlich erkannten Situation umgegangen werden soll, bzw. kann Maßnahmen anstoßen, wie z. B. die Verfolgung einer Person mit statischen oder mobilen Kameras. Während des Einsatzes werden auf den verschiedensten Ebenen automatisiert Entscheidungen getroffen. Nach europäischem Recht sollen Menschen jedoch keiner Beeinträchtigung durch automatisierte Entscheidungen unterworfen sein:

Schutz vor Beeinträchtigung durch Maßnahmen basierend auf automatisierten Entscheidungen

Im Zuge der beabsichtigten EU-Datenschutzreform schlug die Europäische Kommission am 25. Januar 2012 eine neue Datenschutzregelung vor.[248] Sowohl für *freien* (Artikel 20) als auch *behördlichen* Datenverkehr (Artikel 9) werden automatisierte Entscheidungen geregelt.

> *Artikel 20 – Auf Profiling basierende Maßnahmen*
> Eine natürliche Person hat das Recht, nicht einer auf einer rein automatisierten Verarbeitung von Daten basierenden Maßnahme unterworfen zu werden, die ihr gegenüber rechtliche Wirkungen entfaltet oder sie in maßgeblicher Weise beeinträchtigt und deren Zweck in der Auswertung bestimmter Merkmale ihrer Person oder in der Analyse beziehungsweise Voraussage etwa ihrer beruflichen Leistungsfähigkeit, ihrer

248 Sie soll die Verarbeitung personenbezogener Daten vereinheitlichen und die Datenschutzrichtlinie 95/46/EG von 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr ersetzen. [EU Kommission 2012a].

wirtschaftlichen Situation, ihres Aufenthaltsorts, ihres Gesundheitszustands, ihrer persönlichen Vorlieben, ihrer Zuverlässigkeit oder ihres Verhaltens besteht.

Artikel 9 Absatz 1 – Auf Profiling und automatischer Datenverarbeitung basierende Maßnahmen
Die Mitgliedstaaten legen fest, dass Maßnahmen, die eine nachteilige Rechtsfolge für die betroffene Person haben oder sie erheblich beeinträchtigen und die ausschließlich aufgrund einer automatisierten Verarbeitung von personenbezogenen Daten zum Zwecke der Bewertung einzelner Aspekte ihrer Person ergehen, verboten sind, es sei denn, dies ist durch ein Gesetz erlaubt, das Garantien zur Wahrung der berechtigten Interessen der betroffenen Person festlegt.

Die Artikel sind vergleichbar mit Artikel 6a des Bundesdatenschutzgesetzes. Die Variante der EU wurde hier gewählt, da in ihr nicht wie im BDSG der Begriff *Entscheidungen*, sondern *Maßnahmen* gebraucht wird. Dieser Begriff erscheint im Zusammenhang mit Videoüberwachung zutreffender.

OperateurInnen als Legitimation umfangreicher Automatisierung
Je umfangreicher die Videoüberwachung automatisiert wird, desto höher ist die Wahrscheinlichkeit, dass dieses Prinzip verletzt wird. In der Argumentation, die Überwachungsaufgabe dennoch möglichst automatisiert zu gestalten, dient die Rolle der OperateurInnen häufig als Legitimation. In einem Interview äußerte Prof. Andrzej Dziech, Projektkoordinator des *INDECT*-Projektes:

> Die verantwortliche Person analysiert die angezeigten Informationen, urteilt ob es sich nicht nur um einen falschen Alarm handelt und bewertet das Bedrohungslevel. Letztlich, wenn

> es gerechtfertigte Gründe gibt, werden Entscheidungen über relevante Sicherheitsmaßnahmen getroffen. (eigene Übersetzung)[249]

Rossnagel et al. behaupten, mit dem *Drei-Stufen-Modell* könne „sichergestellt werden", dass „die aus der automatisierten Analyse der aufgenommenen Bilder gewonnenen Erkenntnisse selbst nicht unmittelbar zu einer automatisierten Entscheidung führen, sondern die automatisiert gewonnenen Erkenntnisse eine Entscheidungshilfe für die abschließende Entscheidung einer Person mit echtem Entscheidungsspielraum darstellen."[250] Die automatisierte Analyse würde der Letztentscheidung ausschließlich „dienen", „da der Beobachter die Meldung des Systems, dass eine eindeutige Situation vorliegt, inhaltlich auf ihre Stimmigkeit mit den Gesamtumständen der überwachten Situation überprüfen muss, bevor er entscheidet, welche weitergehenden Maßnahmen eingeleitet werden." Sie bewirke aber keine rechtliche Folge, da es lediglich den Beobachter auf die ungewöhnliche Situation aufmerksam mache und ihm diese „optimal präsentiert". Eine „erhebliche Beeinträchtigung" werde im Regelfall erst durch die Gefahrenabwehrmaßnahmen, nicht durch eine weitere und präzisere Beobachtung der Situation erzielt.

Im Folgenden wird argumentiert, dass OperateurInnen dieser Verantwortung nicht gerecht werden können. Den OperateurInnen stehen nur die von Ω gefilterten, vorverarbeiteten und bewerteten Informationen zur Verfügung. Diese reichen für eine „verständige und besonnene"[251] bzw. objektive Lagebeurteilung nicht aus (Kapitel 5.2.1). Außerdem ist durch den Einsatz von Assistenz mit dem psychologischen Effekt eines übersteigerten Vertrauens zu rechnen (Kapitel 5.2.2). Da der Darstellung und

249 **Rutz**, *Interview mit INDECT-Projekt-Koordinator*.

250 **Roßnagel**, **Desoi** ; **Hornung**, „Gestufte Kontrolle bei Videoüberwachungsanlagen : Ein Drei-Stufen-Modell als Vorschlag zur grundrechtsschonenden Gestaltung", S. 699.

251 Ebd., S. 695.

Entscheidung von Ω tendenziell gefolgt wird, muss die Rolle der OperateurInnen als Teil des Automatismus der Überwachung behandelt werden. Die „Letztentscheidung" bzw. die Konsequenzen für die Betroffenen sind dadurch maßgeblich von Ω geprägt. Es muss daher von der Möglichkeit erheblicher Einschränkungen für Betroffene durch Automatisierung ausgegangen werden.
Unberücksichtigt bleibt in dieser Argumentation, dass allein schon die Tatsache der Überwachung an sich eine Beeinträchtigung der Betroffenen darstellt. Neben dem bereits hergeleiteten erhöhten Eingriff in das Recht auf informationelle Selbstbestimmung kommt es zu individuellen und gesamtgesellschaftlichen Auswirkungen. Diese werden anschließend diskutiert.

5.2.1 Mündigkeit und Informiertheit

Wie eben beschrieben, wird den OperateurInnen eine Verantwortung übertragen, deren Wahrnehmung eine umfangreiche Automatisierung der Videoüberwachung legitimieren soll. Der Verantwortung können OperateurInnen nur dann gerecht werden, wenn sie in die Lage versetzt werden, mündige Entscheidungen treffen zu können. Nach Adorno ist derjenige mündig, „der für sich selbst spricht, weil er für sich selbst gedacht hat und nicht bloß nachredet [...]."[252] Eine mündige Entscheidung von OperateurInnen verlangt daher das selbstständige Nachdenken und das eigenständige Bewerten und Abwägen von Informationen.

Veränderung der Aufgabe der OperateurInnen

Bei der Einschätzung, ob dies der Fall sein kann, kommt zum Tragen, wie sich die Aufgabe der OperateurInnen zwischen *manueller* und automatisierter Videoüberwachung unterscheidet. Im Gegensatz zur Generation II, bei der die Assistenz durch das System ausschließlich in der Auswahl

252 **Adorno**, *Gesammelte Schriften : Kulturkritik und Gesellschaft II*, S. 785.

der Videobilder besteht, übernimmt Ω auch weitere vormals kognitive Aufgaben der OperateurInnen. Mit Ω wird versucht, den OperateurInnen sowohl die mentale Filterung (Auswahl der näher zu beobachtenden Geschehnisse) als auch die Rekontextualisierung weitestgehend abzunehmen. Ω entscheidet über Risikoorte, Risikopersonen und Risikoobjekte und wählt unter den aufgenommenen Bilder aus. Außerdem berechnet Ω zusätzliche, wertende Informationen und präsentiert diese und die Bilder in aufbereiteter Form den OperateurInnen. Außerdem impliziert Ω direkt über konkrete Empfehlungen oder indirekt über die dargestellten Informationen, wie auf die vermeintlich klassifizierte Situation reagiert werden sollte. Die Aufgabe der OperateurInnen besteht dann darin, die Angaben des Systems zu überprüfen, einen Alarm oder eine explizite oder indirekte Handlungsempfehlung zu bestätigen oder sie zu ignorieren.

Ein Beispiel einer Empfehlung bei dem *Drei-Stufen-Modell* ist der Vorschlag, in eine höhere Stufe zu wechseln, eine bestimmte Person näher zu verfolgen oder die wie auch immer gearteten Maßnahmen für eine von Ω verdächtigte Person einzuleiten.

Voraussetzung für mündige Entscheidungen beim Einsatz von Ω

Um als OperateurIn von Ω eine mündige Entscheidung treffen zu können, muss 1) eine mögliche Diskrepanz zwischen Gezeigtem und der Realität *bewusst* sein, 2) nachvollzogen werden können, *wie* die Empfehlung oder die Informationen die eine gewisse Handlung implizierte zustande kam, und 3) müssen die Informationen mit dem *eigenen Verständnis der Geschehnisse* abgeglichen werden.

Sensorische und semantische Kluft

Das Mehr an Informationen, die den OperateurInnen präsentiert werden, scheint gegenüber *manueller Videoüberwachung* eine größere Informiertheit zu bewirken. Es kann jedoch weder der überwachte Raum noch

die in ihm stattfindenden Geschehnisse akkurat dargestellt werden. Dies liegt zum Einen an dem bereits beschriebenen technischen Filter und den mehrdeutigen Interpretationsmöglichkeiten der Bilder. Zum Anderen liegt der Grund in der Funktionsweise der in Kapitel 3 und besonders in Kapitel 3.1 beschriebenen Techniken zur automatisierten Interpretation von Verhalten. Auf den unteren Abstraktionsebenen – vor allem der Bildverarbeitung – kommt es zu Ungenauigkeiten bei der quantitativen Erhebung der Daten. Diese wird als *sensorische Kluft (sensoric gap)*[253] bezeichnet. Auf höheren Abstraktionsebenen – vor allem beim Einsatz von künstlicher Intelligenz – entsteht durch die Uneindeutigkeit der Interpretation der Informationen eine *semantische Kluft* (*semantic gap*). Eine Informations- und Komplexitätsreduktion ist für die algorithmische Verarbeitung einerseits zwingend erforderlich. Andererseits folgt aus der Algorithmisierung wiederum auch eine Komplexitätsreduktion. Durch Modellierung, statistische Auswertung und Verarbeitung von Wahrscheinlichkeiten, die zentrale Konzepte der beschriebenen Techniken sind, gehen Informationen verloren, verlieren an Aussagekraft und werden komplexe Sachverhalte wie Verhalten stark vereinfacht. Trotz dieser Defizite besteht eine weit verbreitete Erwartung, dass die Leistungsfähigkeit des visuellen Teils eines Systems wie Ω mit der eines Menschen verglichen werden kann.[254] Dies resultiert aus einer zu starken Vereinfachung der Leistungsfähigkeit des Menschen und führt zu einer Überschätzung der Automatisierung.

Überschätzung

Erwartungen an die Zuverlässigkeit und Fehlerfreiheit technischer Systeme sind sehr hoch. Von Technik und computerisiert berechneten Ergeb-

253 **Baiget** et al., „Observing Human Behavior in Image Sequences: the Video-Hermeneutics Challenge", S. 1.

254 **Musik**, „The thinking eye is only half the story: high-level semantic video surveillance", S. 340.

nissen wird meist angenommen sie seien objektiv und wertfrei.[255] Außerdem wird Automatisierung in vielen Fällen mehr Vertrauen entgegen gebracht als einem Menschen, der die gleiche Aufgabe ausführt.[256] Auf Beides wird in den folgenden Kapiteln näher eingegangen. Auch die Europäische Kommission erkennt das Problem, dass Automatisierung von Entscheidung zu automatischer Annahme von Validität und begleitend eine Reduzierung der Prüfung und Verantwortung von Menschen führen kann:

> The result produced by the machine, using more and more sophisticated software, and even expert systems, has an apparently objective and incontrovertible character to which a human decision-maker may attach too much weight, thus abdicating his own responsibilities.[257]

Im Rahmen der Usabilityforschung wurde erkannt, dass auf die Glaubwürdigkeit und Interpretation selbst subtile Aspekte der Darstellung signifikante Auswirkungen haben können und maßgeblichen Einfluss auf die resultierenden Entscheidungen und Handlungen der AnwenderInnen haben. Jared Spool schildert zum Beispiel einen eindrucksvollen Fall, bei dem nur die Beschriftung eines Buttons das Kaufverhalten in einem Internetshop vollkommen veränderte.[258]) Von entscheidender Bedeutung beim Einsatz von Ω ist also auch die gewählte Darstellung der Informationen.

---

255 **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems", S. 195.

256 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 19.

257 **Bygrave**, „Minding the machine: Article 15 of the EC Data protection Directive and Automated Profiling", S. 25.

258 **Spool**, *The $300 Million Button*.

Illusion von Vollständigkeit und Objektivität

In Ω werden Kamerabilder in eine dreidimensionale virtuelle Realität eingebettet. In diese werden die Resultate der Analysen auf verschiedenste Weise eingeblendet. Die vorhandene sensorische und semantische Kluft wird dadurch verschleiert. Die geschaffene Realität wirkt so umfassend oder der Realität des überwachten Raumes in ihrem visuellen Erscheinungsbild so ähnlich, dass den OperateurInnen nicht nur der Eindruck von Korrektheit und Objektivität der Informationen, sondern auch der Eindruck eines vollständigen Situationsbewusstsein vermittelt wird. Bei *manueller Videoüberwachung* ist das Fehlen von Informationen offensichtlicher. Dieser Informationsmangel muss einkalkuliert werden und die sichtbaren Informationen vorrangig mit eigenem Wissen und Erfahrungen über Menschen und den Raum rekontextualisiert werden. Durch die vermeintliche Vollständigkeit scheint die Notwendigkeit einer Rekontextualisierung beim Einsatz von Ω reduziert oder sogar aufgehoben.

Qualifizierung und Ausbildung der OperateurInnen

Was ein Mensch bei *manueller Videoüberwachung* mit Wissen und Erfahrung über den überwachten Raum ausgleichen musste, übernimmt bei automatisierter Videoüberwachung je nach Konfiguration in bestimmtem Umfang das System. Der (vermeintlich) geringere Bedarf von Qualifizierung für diese Aufgabe ist im Interesse der Betreiber, da die Senkung der Personalkosten Ziel *manueller Videoüberwachung* war[259] und auch der Automatisierung von Videoüberwachung ist. Vor diesem Hintergrund ist damit zu rechnen, dass Betreiber mitunter Menschen einsetzen werden, denen eine angemessene Ausbildung nicht abverlangt bzw. zugestanden wird. Schon jetzt bedarf das gewerbliche Bewachen fremden Lebens oder Eigentums einer behördlichen Erlaubnis, die bereits nach 40 Lehrstun-

259 **Töpfer**, „Videoüberwachung als Kriminalprävention? Plädoyer für einen Blickwechsel", S. 275.

den[260] erworben werden kann.[261] Es muss davon ausgegangen werden, dass für Ω eher selten gut ausgebildete Personen als OperateurInnen eingesetzt werden. Durch Vernetzung der Anlagen und die Möglichkeit der Verlagerung der Überwachungsaufgabe besteht sogar die erhöhte Wahrscheinlichkeit, dass Personen die Aufgabe übernehmen, die keinerlei oder nur geringe Kenntnisse über den überwachten Raum haben. In diesem Falle ist der Einfluss von Ω auf die Sichtweise der OperateurInnen auf die Geschehnisse um so größer. Doch auch gut ausgebildeten OperateurInnen kann mit dem Maß an Informationen eine angemessene Einschätzung nicht immer abverlangt werden.

Darstellbarkeit der automatischen Entscheidungsfindung

Bei Ω handelt es sich um ein System, das eine Fülle von Informationen auf eine komplexe Weise verarbeitet. Um eine Entscheidung oder Empfehlung von Ω mündig zu bestätigen oder abzulehnen, müsste der Entstehungsprozess nachvollziehbar dargestellt werden. Die vielschichtigen, miteinander verwobenen, stark technisch bedingten Abläufe für OperateurInnen in der zu Verfügung stehenden Zeit verständlich darzustellen, scheint jedoch eine unlösbare Aufgabe zu sein. Dies liegt nicht vornehmlich an intellektuellen Grenzen der OperateurInnen, sondern vor allem an der Darstellbarkeit der Funktionsweise einzelner Verarbeitungsschritte. Beispielsweise ist selbst für EntwicklerInnen meist nicht nachvollziehbar, welche Werte eines neuronalen Netzes in der Anwendung welche Folgen hat. Ein weiteres Beispiel ist die Visualisierbarkeit des Abgleichs einer erfassten Gesichtsbewegung mit einer anhand tausender Gesichter trainierten Mimikdatenbank zu visualisieren. Tatsächlich werden bei au-

260 Vgl. Emagister.de, *Kurs Unterrichtung für Bewachungspersonal und Gewerbetreibende nach § 34a GewO* Internet: `http://www.emagister.de/kurse_34a_gewo-esn300855.htm`.

261 Die Erlaubnis genannt „§ 34a-Schein" wird nach einer *Sachkundeprüfung* gemäß § 34a der deutschen Gewerbeordnung erteilt.

tomatisierter Mimikerkennung über 10.000 mögliche Gesichtsausdrücke unterschieden.[262]

Entwurf eines eigenen Bildes der Situation

Alle Informationen werden vom System ausgewählt, kombiniert, bewertet und allein auf dieser Grundlage dargestellt. Die nach der Filterung übrig gebliebenen Informationen haben unweigerlich eine von Ω signifikant beeinflusste Konnotation. Viel mehr noch als durch den *technischen Filter* werden hier also der betrachtete Raum und in besonderem Maße auch die Betrachtungsweise a priori festgelegt. Die Gegebenheit der dritten Voraussetzung für eine mündige Entscheidung, nämlich der Abgleich des Gezeigten mit der eigenen Wahrnehmung, ist dementsprechend ebenfalls anzuzweifeln. Den OperateurInnen wird gar nicht die Möglichkeit gegeben, eine eigene Interpretation der Geschehnisse zu entwerfen. Durch das Nutzen zusätzlicher Datenquellen wie Personendatenbanken oder Analyseergebnissen von sozialen Netzwerken, die sich dem Zugriff der OperateurInnen (z. B. aus Datenschutzgründen) vollends entziehen, wird die Wahrscheinlichkeit einer Nachvollziehbarkeit und eines Gegenentwurfs weiter verringert. In der Praxis führt der Vorteil der Automatisierung, eine große Menge an Bildern und Daten verarbeiten zu können, zum Nachteil der Überschaubarkeit. Mit automatisierter Videoüberwachung hat ein Mensch höchstwahrscheinlich für signifikant mehr zu überwachenden Raum die Verantwortung, als es noch der Fall bei *manueller Videoüberwachung* war.[263] So ist anzunehmen, dass zusätzlich auch aus der Limitierung der Kapazitäten der OperateurInnen kein Gegenentwurf zur dargestellten Situation entwickelt werden kann.

---

262 **Musik**, „The thinking eye is only half the story: high-level semantic video surveillance", S. 345.

263 Visuelle Überlastung (Visual Overload) wird auch von Suarez als Ursache angeführt, dass bei automatischen Kampfdrohnen die Entscheidung weiter vom Menschen wegrückt.

Der Faktor Zeit

Zu den bereits genannten Schwierigkeiten der OperateurInnen, eine mündige Entscheidung zu treffen, kommt der Zeitfaktor. Liefert Ω eine Interpretation, die für eine Gefahr spricht, so steht den OperateurInnen für viele denkbare Szenarien keine Zeit für Verifikation zur Verfügung, weil Gefahr in Verzug scheint. Die Kosten, die ein Operateur durch einen bestätigten Fehlalarm zu tragen hätte, bzw. die Kosten, die dadurch für Betroffene entstünden, wiegen bei weitem nicht die Kosten auf, die bei menschengeschuldetem *miss* für Betroffene, aber auch für OperateurInnen selbst entstehen würden. OperateurInnen werden nach dieser Kosten-Nutzen-Rechnung, ähnlich wie Ω selbst, dazu tendieren, „empfindlich" zu sein, Interpretationen von Ω tendenziell zu akzeptieren und Alarmen und Empfehlungen dementsprechend Folge zu leisten.

5.2.2 Übersteigertes Vertrauen in Automation

Bei der bisherigen Argumentation, warum nicht von mündigen Entscheidungen der OperateurInnen auszugehen ist, stand die Quantität und Qualität der Informationen im Mittelpunkt, die zur Prüfung der expliziten und impliziten Empfehlungen zur Verfügung gestellt werden. Im Folgenden wird nun die Perspektive der Mensch-Technik-Interaktion eingenommen und die Qualität der *Überprüfung* dieser Informationen durch die OperateurInnen betrachtet. Diese Qualität wird entscheidend vom Vertrauen in die Automatisierung und die Assistenz durch das System bestimmt.

Untersuchung unerwünschter Effekte von Automation

Die Mensch-Technik-Interaktion wird von der eng mit der Informatik verknüpften Ingenieurpsychologie untersucht. Ihr Ziel ist es, die Gestaltung technischer Systeme nach den Bedürfnissen der sie benutzenden Menschen auszurichten. Eins der Ziele ist es, Probleme zu identifizie-

ren, deren Ursachen zu ermitteln und mögliche Gegenmaßnahmen vorzuschlagen. Es wurden verschiedene Versuche einer Strukturierung unerwünschter Effekte von Automation unternommen. Neben der Abnahme von Situationsbewusstsein sogenannter *out-of-the-loop-unfamiliarity* und dem Verlust von Fähigkeiten, auf die hier nicht weiter eingegangen werden kann, wurde dabei stets der Problembereich eines *übersteigerten Vertrauens* identifiziert.[264] Die Adäquatheit des Verhältnisses von *Vertrauen*, das einem System entgegengebracht wird, und der *Zuverlässigkeit*, die es tatsächlich erbringen kann, ist für die Gestaltung von Technik von großem Interesse: Ein zu geringes Maß an Vertrauen ist unerwünscht, da dies zu einer mangelnden Nutzung führt. Ein übersteigertes Vertrauen hingegen führt zu Nachlässigkeit und mangelnder Kontrolle. Hier soll nur das übersteigerte Vertrauen betrachtet werden.

Übersteigertes Vertrauen in Assistenz

Mosier, Skita und Burdick gehen davon aus, dass die Einführung von Assistenzsystemen immer das Risiko birgt, dass eine eigene aktive Entscheidungsfindung durch Automation ersetzt wird.[265] Das Ergebnis einer Reihe von Studien war, dass trotz Hinweis auf die Fehleranfälligkeit des Testsystems, extrem hohe Fehlerraten (bis zu 55% *omission* Fehler[266] und sogar 100% *commission* Fehler) bei den Versuchspersonen beobachtet wurden. Es muss daher bei der Einführung von Assistenz von der Möglichkeit eines *erhöhten* Risikos von automationsbezogenen Fehlern ausgegangen werden.

Im Folgenden werden *complacency* und *automation bias* vorgestellt – zwei Phänomene der Mensch-Technik-Interaktion, deren Basis ein übersteiger-

264 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 12.

265 **Mosier**, **Skitka** ; **Burdick**, „Automation bias: Decision making and performance in high-tech cockpits".

266 Die Begriffe *omission* und *commission* Fehler werden weiter unten eingeführt.

tes Vertrauen in Automation ist. Beim Einsatz von Ω würde übersteigertes Vertrauen in die Automation und die Assistenz durch das System dazu führen, dass Empfehlungen und Interpretationen des Systems akzeptiert werden, ohne dass verifizierende oder falsifizierende Informationen ausreichend berücksichtigt werden und dass Gefahrensituationen, die nicht von Ω erkannt werden (*miss*) auch von den OperateurInnen übersehen werden. Nach der Beschreibung eines Experimentes, in dem *complacency* über das Informationssuchverhalten operationalisiert nachgewiesen wurde, wird der Versuchsaufbau mit Ω verglichen. Damit soll gezeigt werden, dass die Phänomene auch beim Einsatz von Ω zu erwarten sind und daher auch aus psychologischen Gründen die Mündigkeit der Entscheidungen der OperateurInnen in Frage gestellt werden muss und diese nicht als Legitimation einer umfangreichen Automatisierung angeführt werden können.

Automation erfordert Vertrauen

Jennifer Bahner integriert die vormals separat betrachteten Konzepte *complacency* und *automation bias* und überträgt sie vom Luftfahrtkontext auf den Einsatz von Assistenzsystemen im allgemeineren Anwendungskontext der Prozesssteuerung.[267] *Automation* wird in diesem Zusammenhang begriffen als die Ausführung von Tätigkeiten durch eine Maschine, die ein Mensch ausführen könnte. Sie dient der Reduzierung der Beanspruchung und des Trainingsaufwandes, der Vermeidung von Fehlern der OperateurInnen, sowie der Steigerung der Verlässlichkeit und Wirtschaftlichkeit des Systems.[268] Während Automatisierung vormals vor allem für manuelle Regelungs- und Steuerungsfunktionen angestrebt wurde, werden immer mehr auch kognitive Funktionen der Urteils- und Entscheidungsfindung dem technischen System übertragen. Die daraus resultierende Kom-

267 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias".

268 Ebd., S. 11.

plexität kann für den Menschen nur dann handhabbar gemacht werden, wenn ein hohes Maß an Vorverarbeitung und Integration der Daten vorgenommen wird.

Vertrauen

Diese Komplexitätsreduzierung verlangt den BenutzerInnen ein bestimmtes Maß an Vertrauen ab. *Vertrauen* wird verstanden als „die Überzeugung einer Person, dass ihr Interaktionspartner (Mensch oder Automation) sie bei der Erreichung ihrer Ziele in einer durch Unsicherheit gekennzeichneten Situation unterstützen wird".[269] Je nach Vertrauen und Vertrauenswürdigkeit (hier Leistungsfähigkeit) kann das subjektive Vertrauen als *mangelnd*, *angemessen* oder *übersteigert* beschrieben werden. Ein übersteigertes Vertrauen führt zu *abuse*, dem unangemessenen Einsatz und zu *misuse*, der Überschätzung durch die AnwenderInnen.[270]
In vielen Fällen ist das Vertrauen in Automation von vornherein größer als in Menschen, die die gleiche Aufgabe erfüllen. Gleichzeitig ist das Vertrauen in Automation jedoch auch anfälliger.[271] Ein Fehler in einem Modul kann sich auch auf das Vertrauen in andere funktional eigentlich unabhängige Module auswirken und so zu einem umfassenden Vertrauensentzug führen.[272] Entscheidende Faktoren, die das Vertrauen beeinflussen, sind potentielle Kosten durch Fehler, die Nachvollziehbarkeit und Vorhersagbarkeit von Automationsfehlern und das Selbstvertrauen der OperateurInnen in die eigene manuelle Ausführung. Daraus ergibt sich, dass mangelndem Vertrauen mit Nachvollziehbarkeit, Zuverlässigkeit bzw. hoher *Empfindlichkeit* bei kostenintensiven Fehlern entgegengewirkt werden kann. Durch derartige Maßnahmen und Eigenschaften erhöht sich jedoch das ohnehin erhöhte Risiko eines übersteigerten Ver-

269 Ebd., S. 21.
270 Ebd., S. 14.
271 Ebd., S. 19.
272 Ebd., S. 22.

trauens. Dieses führt nicht nur zu reduzierter Kontrolle und Verifikation, sondern bewirkt zusätzlich, dass Informationen, die Zweifel an der Korrektheit der Automation aufkommen lassen müssten, zwar registriert aber nicht zur Falsifikation genutzt werden.[273] Die Phänomene *complacency* und *automation bias* werden im Folgenden vorgestellt.

Complacency

Der Begriff *complacency* wurde geprägt im Zusammenhang mit der zunehmenden Automatisierung im Luftfahrtbereich – der Effekt ist aber in allen Bereichen, die durch Automatisierung gekennzeichnet sind, zu erwarten.[274] *Complacency* ist ein Merkmal der Mensch-Technik-Interaktion, das bei kontinuierlich überwachender Kontrolle automatisierter Systeme zur Erkennung von Automationsfehlern zu beobachten ist. Es äußert sich durch ein übersteigertes Vertrauen in die Funktions- und Leistungsfähigkeit der Automation. Dieses kann zu unzureichender Überwachung des Systems führen, aus der ein Übersehen kritischer Systemzustände folgen kann. *Complacency* wird daher als einer von mehreren Determinanten für den Verlust von Situationsbewusstsein aufgefasst. Es lassen sich drei Faktoren unterscheiden, die zur Entstehung beitragen: Merkmale der *Automation*, der *Person* und des *situativen Kontextes*. Neben den Faktoren, die das Vertrauen in Automation erhöhen, wird *complacency* maßgeblich durch eine hochreliable Automation begünstigt. Zu den begünstigenden Personenmerkmalen gehören die Neigung zu Langeweile, zu kognitiven Fehlern und eine geringe Selbstwirksamkeitserwartung. Auch die Sorglosigkeit sowie die Risiko- und Ungewissheitstoleranz tragen dazu bei.[275] Zusammen ergeben die Merkmale eine Verhaltenstendenz. Die bisher genannten sind notwendige, aber wie vermutet wird, nicht hinreichende Be-

273 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 27.
274 Ebd., S. 27 f.
275 Ebd., S. 34.

dingungen für *complacency*. Eine moderierende Rolle übernimmt die Eigenschaft der Situation. Wesentliche Grundvoraussetzung ist das Vorhandensein konkurrierender Zielstellungen. Es wird dabei tendenziell diejenige Aufgabe der Automation überlassen, deren korrekter Bearbeitung stärker vertraut wird. Ein weiterer situativer Aspekt ist das wahrgenommene Risiko. Eine übermäßige Reduktion der Überwachungsintensität wird vor allem dann vorgenommen, wenn das Risiko als gering eingestuft wird. Bahner nimmt zusätzlich an, dass es zwischen Verhaltenstendenz und der Verhaltenskonsequenz eine Rückkopplung gibt: Machen die OperateurInnen trotz mangelnder Überwachung nicht die Erfahrung negativer Konsequenzen, kann dies zu einem Lernprozess führen, der die Verhaltenstendenz zu übersteigertem Vertrauen verstärkt.[276] Das genaue Zusammenwirken der drei Faktoren ist bisher nur unzureichend untersucht worden.

Automation bias

Während *complacency* besonders im Kontext kontinuierlicher Überwachung automatisierter Systeme beschrieben wurde, wird der Begriff *automation bias* im Zusammenhang mit dem Nutzen von Assistenzsystemen gebraucht. Mosier und Skitka verstehen unter *automation bias* die Tendenz, automatisierte Hinweise als heuristischen Ersatz für wachsames Suchen und Verarbeiten von Informationen zu nutzen.[277] Es werden zwei mögliche Fehlertypen unterschieden: *omission* und *commission* Fehler. *Omission* Fehler werden begangen, wenn ein vorliegender kritischer Systemzustand vom Assistenzsystem nicht angezeigt wird (*miss*) und auch der Operateur oder die Operateurin ihn übersieht. Ein *commission* Feh-

276 **Manzey** ; **Bahner**, „Vertrauen in Automation als Aspekt der Verlässlichkeit von Mensch-Maschine-Systemen".

277 **Mosier** ; **Skitka**, „Human Decision Makers and Automated Decision Aids: Made for Each Other?", S. 205.

ler liegt vor, wenn einer falschen Direktive des Assistenzsystems Folge geleistet wird, z. B. bei einem *Fehlalarm*.

Als Ursache für *commission* Fehler wird das Unterlassen des Suchens nach verifizierenden oder falsifizierenden Informationen (Informationssuchverhalten) angesehen. Entweder wird der Abruf verfügbarer Informationen zur Prüfung des Hinweises ganz unterlassen oder die Informationen verzerrt verarbeitet. Es werden drei Verzerrungsmechanismen unterschieden: *assimilation bias*, *confirmational bias* und *discounting bias*. Von *assimilation bias* spricht man, wenn mehrdeutige Informationen vorliegen, jedoch ausschließlich konsistent zum Systemhinweis interpretiert werden. *Confirmational bias* besteht im Ausblenden aller zum Systemhinweis inkonsistenten Informationen. Beim *discounting bias* werden widersprüchliche Informationen zwar wahrgenommen, jedoch in ihrer Bedeutung herabgesetzt und nicht für die eigene Entscheidung berücksichtigt.

Integration von *complacency* und *automation bias*

Grund für *omission* und *commission* Fehler ist neben *Vigilanzproblemen* und Fokussierung auf eine *konkurrierende Aufgabe* auch *complacency*.[278] *Complacency* führt durch mangelnde *Überwachung* zu *omission* Fehlern, und durch mangelnde *Überprüfung* zu *commission* Fehlern (Abb. 24).[279] Sowohl *complacency* als auch *automation bias* können also auf ein übersteigertes Vertrauen in Automation zurückgeführt werden.

Vermeidbarkeit

Bahner bespricht Möglichkeiten zur Vermeidung bzw. Reduzierung von *complacency*.[280] Es kann davon ausgegangen werden, dass sich ein ausgeprägtes *Verantwortungsgefühl* der OperateurInnen für ihre Aufgabe posi-

278 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 50.
279 Ebd., S. 51.
280 Ebd.

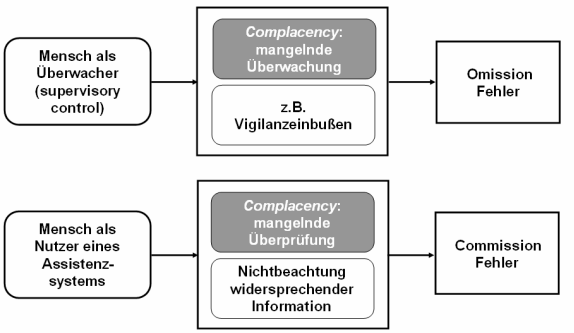


Abb. 24: *Complacency* im Kontext von Systemüberwachung und Nutzung von Assistenzsystemen.

tiv auf die Fehlerrate auswirkt. Ein möglicher Ansatzpunkt zur Verbesserung der Fehlerraten schien daher bis dato darin zu liegen, den OperateurInnen ihre Verantwortung deutlich zu machen. Es wird jedoch angenommen, dass es sich dabei um ein interindividuell unterschiedlich stark ausgeprägtes Personenmerkmal handelt, das nur sehr begrenzt induziert werden kann.[281]

Auch die Vermutung, dass die Durchführung der Überwachungsaufgabe mittels *Zusammenarbeit* mehrerer Personen die Fehlerraten positiv verändern würde, konnte nicht bestätigt werden. Bei der räumlich getrennten Durchführung konnte entgegen der Erwartungen sogar ein Anstieg der Fehlerrate festgestellt werden. Es wird angenommen, dass dieser Effekt mit der Abgabe und Reduzierung der Verantwortung an eine zweite Person einhergeht. Von einer Möglichkeit der Verbesserung des Informationssuchverhaltens durch Teamarbeit kann daher nicht ausgegangen werden.

In früheren Experimenten wurde untersucht, inwiefern ein explizites *Trai-*

281 Ebd., S. 45.

ning gegen *automation bias* die Fehlerraten verbesserte. Ergebnis war, dass *commission* und *omission* Fehler relativ robuste Phänomene sind. Weder durch vorhergehende Instruktionen, dass Assistenzsystem zu überprüfen, noch durch Aufforderung während der Ausführung der Aufgabe, wurden die Fehlerraten beeinflusst.[282]

Experiment zur Untersuchung der Phänomene

Bahner nutzte im Rahmen eines Experimentes eine adäquatere Operationalisierung von *complacency*, als bei vorangegangenen Experimenten, in dem nun das Informationssuchverhalten berücksichtigt wurde. Probanden hatten die Aufgabe, das Lebenserhaltungssystem einer virtuellen Raumstation zu überwachen, Fehler zu diagnostizieren und gegebenenfalls manuell zu beheben. Werte von Subsystemen wie Sauerstoff, Druck, Kohlenstoffdioxid, Temperatur und Ähnlichem sollten innerhalb eines bestimmten Bereiches liegen. Bei der Diagnose, Behebung, Fehlerdetektion, -diagnose und -management wurden sie von einem Assistenzsystem, bestehend aus zwei Teilen unterstützt – einer *Alarmfunktion* und einer *Diagnosefunktion*. Nach einem unspezifischen Masteralarm der *Alarmfunktion* bei Abweichung vom gewünschten Zustand, wurden von der *Diagnosefunktion* bestimmte Handlungsschritte empfohlen. Zusätzlich hatten die ProbandInnen einfache, konkurrierende Aufgaben zu erledigen. Alle Informationen und Bedienelemente waren auf einem Bildschirm schematisch oder als Werte dargestellt. Im Verlauf kam es zu verschiedenen unerwünschten Zuständen des Lebenserhaltungssystem, die je nach Werten der Subsysteme, nach bestimmten vorher bekannten Regeln behandelt werden mussten. In zwei separaten Experimenten wurde zum Einen *complacency* gegenüber der *Alarmfunktion* und zum Anderen gegenüber der *Diagnosefunktion* gemessen. In einem Durchlauf kam es zum

282 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 46.

nicht angezeigten Ausfall der *Alarmfunktion*, im Zweiten zu Fehlern der *Diagnosefunktion* und somit zu falschen Empfehlungen. Ermittelt wurde das Informationssuchverhalten in fehlerfreien Phasen, sowie die Raten von *commission* und *omission* Fehlern unter Berücksichtigung der Frage, welche Informationen für die Wahl der Behebungsmethode aktiv hätten abgerufen werden müssen.

Ergebnis des Experiments

Das Ergebnis zeigte, dass Fehlererfahrungen der *Alarmfunktion* bzw. der *Diagnosefunktion* sich *spezifisch* auf das Vertrauen in die jeweilige Funktion auswirken. Spezifische *complacency* Effekte konnten durch Training mit Fehlererfahrungen zwar signifikant reduziert, jedoch nicht gänzlich vermieden werden. Es wurde deutlich, dass auch nach Fehlererfahrungen noch *commission* Fehler auftreten. Dies ist sogar dann der Fall, wenn falsifizierende Informationen zwar abgerufen aber nicht beachtet werden. Alle drei oben genannten Verzerrungsmechanismen wurden beobachtet. Zusätzlich wurden von den Versuchspersonen sogar vermeintlich verifizierende Informationen erinnert, die so vom System nie gegeben wurden.

Schlussfolgerung für Gestaltung von Trainings

Aus den Ergebnissen konnten Schlussfolgerungen über die Gestaltung von Trainings abgeleitet werden. Im Gegensatz zur gängigen Praxis dürfen Trainings nicht nur einzelne Automationsfehler adressieren oder auf den Umgang mit Komplettausfällen fokussieren. Es ist ein umfangreiches, mit viel Aufwand gestaltetes Training notwendig. Die Tatsache, dass vorhergehende verbale Instruktionen keinen messbar verbessernden Effekt auf die Fehlerraten hatten, erschwert die Gestaltung des Trainings zusätzlich, da eine theoretische Unterrichtung der OperateurInnen über mögliche Fehler keinen signifikanten Effekt hat.

Übertragbarkeit der Phänomene auf Ω

Bahner konnte die Vermutung empirisch bestätigen, dass *complacency* nicht nur bei überwachender Kontrolle automatisierter Systeme ein ernst zu nehmendes Problem darstellt, sondern auch bei der Nutzung von Assistenzsystemen sehr wahrscheinlich ist. Im beschriebenen Experiment zeigten alle Probanden *complacency* gegenüber der*Diagnosefunktion* des Assistenzsystems, nutzten also nicht alle zur Überprüfung notwendigen Informationen aus.[283]

Zur Veranschaulichung, dass die beim Experiment beobachteten Phänomene auch beim Einsatz von Ω mit hoher Wahrscheinlichkeit auftreten werden, sollen die Voraussetzungen der Entstehung von *complacency* betrachtet und Ω mit dem Versuchsaufbau des Experiments verglichen werden.

Voraussetzungen für das Entstehen

Zu den personenspezifischen Voraussetzungen für *complacency* wie Neigung zur Langeweile und Selbstwirksamkeitserwartung der OperateurInnen kann an dieser Stelle keine Aussage getroffen werden. Auch zur Reliabilität von Ω könnten nur Mutmaßungen aufgestellt werden. Festzustellen ist jedoch, dass im Gegensatz zum Experiment, bei dem alle Fehler direkt anhand der Werte der Subsysteme sichtbar werden, Fehler von Ω nicht unbedingt sichtbar werden und so nicht dazu beitragen können, dass es in bestimmten Aspekten als unverlässlich eingeschätzt wird. Einfluss auf das Vertrauen hat nur die individuell wahrgenommene Reliabilität selbst. Da Fehler nicht unbedingt für die zuständigen OperateurInnen unmittelbar wahrnehmbar sind (z. B. ein „verpasster" Diebstahl), kann dies nach dem Prinzip der Rückkopplung durch Ausbleiben negativer Konsequenzen zu einem Lernprozess führen, der das individu-

283 **Bahner**, „Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias", S. 93.

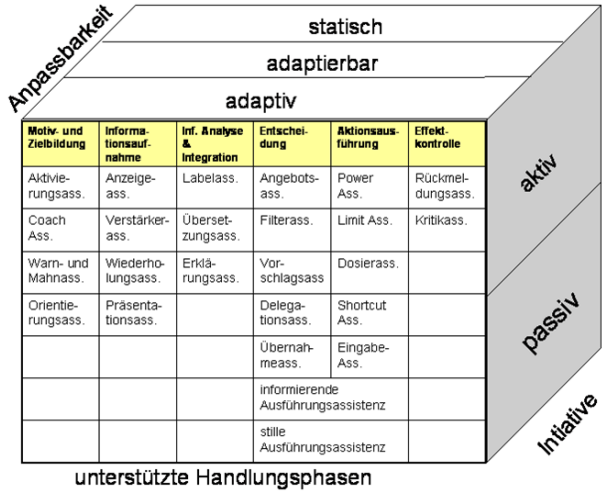
elle Vertrauen verstärkt. *Merkmale der Situation*, die Voraussetzung zur Entstehung von *complacency* sind, sind beim Einsatz von Ω mit hoher Wahrscheinlichkeit gegeben. Konkurrierende Zielstellungen sind zu erwarten, da OperateurInnen in der Regel für einen größeren zu überwachenden Bereich zuständig sind. OperateurInnen sind durch ihre Aufgabe wie in 2.3.1 beschrieben eher kognitiv überlastet als unterfordert. Es ist davon auszugehen, dass OperateurInnen z. B. aus Kostengründen trotz Unterstützung durch Ω auch weiterhin grundsätzlich *mindestens* so viel kognitive Last auferlegt wird, wie ihnen zuzumuten ist.

Vergleich des Versuchsaufbaus mit Ω

Das Lebenserhaltungssystem im Experiment entspricht beim Einsatz von Ω dem überwachten Raum, in dem ein gewünschter Zustand aufrecht erhalten bzw. hergestellt werden soll. Maßnahmen zur Korrektur und Einflussnahme auf das Lebenserhaltungssystem entsprechen den Maßnahmen, die OperateurInnen selbst durchführen können, z. B. eine Lautsprecherdurchsage oder durch Alarmierung des Personals vor Ort. Während das Assistenzsystem im Experiment lediglich aus der *Alarmfunktion* und der *Diagnosefunktion* besteht, ist Ω bei weitem komplexer. Mit Hilfe des Assistenzwürfels (Abb. 25), einer Systematik von Assistenz bei der Mensch-Technik-Interaktion[284], ließen sich eine ganze Reihe weiterer Assistenten identifizieren, die hinsichtlich ihrer Ursache für *complacency* untersucht werden könnten.

Eine umfangreiche Analyse ist an dieser Stelle jedoch nicht nötig, da bereits die Assistenten von Ω, die mit den zwei Assistenten des Experiments (*Alarm-* und *Diagnosefunktion*) direkt vergleichbar sind, zeigen, dass auch beim Einsatz von Ω mit *complacency* und *automation bias* gerechnet werden muss. Die dem Experiment entsprechenden Assistenten von Ω sind

284 **Wandke**, **Wetzenstein** ; **Polkehn**, „Handlungsbezogene Elementarbausteine für Fahrerassistenzsysteme“.



Motiv- und Zielbildung	Informationsaufnahme	Inf. Analyse & Integration	Entscheidung	Aktionsausführung	Effektkontrolle
Aktivierungsass.	Anzeigeass.	Labelass.	Angebotsass.	Power Ass.	Rückmeldungsass.
Coach Ass.	Verstärkerass.	Übersetzungsass.	Filterass.	Limit Ass.	Kritikass.
Warn- und Mahnass.	Wiederholungsass.	Erklärungsass.	Vorschlagsass	Dosierass.	
Orientierungsass.	Präsentationsass.		Delegationsass.	Shortcut Ass.	
			Übernahmeass.	Eingabe-Ass.	
			informierende Ausführungsassistenz		
			stille Ausführungsassistenz		

Abb. 25: Assistenzwürfel – Nach dieser Systematik ließe sich Ω auf weitere Merkmalen von Assistenz hin untersuchen.

die *Alarmfunktion*, die unerwünschtes Verhalten in irgendeiner Art und Weise meldet, und das Informationssystem, das Informationen und Interpretationen darstellt und darauf basierend Maßnahmen explizit oder implizit empfiehlt. Da die Aufgabe der Interpretation der Geschehnisse zur Aufgabe von Ω gemacht wird, besteht die Aufgabe der OperateurInnen darin, die Alarme, Interpretationen und Empfehlungen zu akzeptieren und entsprechende Maßnahmen einzuleiten oder sie zu ignorieren.
Da durch die Assistenz dieser Aufgabe jedoch mit *complacency* gerechnet werden muss, kann von Ω eine Kette von Ereignissen verursacht werden, ohne dass OperateurInnen dies in Folge eigener Denkleistung verhindern. Selbst wenn das technische System allein nicht in der Lage ist, Entscheidungen umzusetzen und z. B. verdächtige Personen festzunehmen, „unschädlich“ zu machen oder andersartig manipulierend in den

überwachten Raum zu wirken, so besteht doch die Möglichkeit, dass allein von Ω Maßnahmen ausgewählt, angestoßen oder zumindest entscheidend geprägt werden, die im Verlauf der Geschehnisse nach einem Alarm bis zur letzten Konsequenz durchgeführt werden und erhebliche Nachteile für die Betroffenen bedeuten können.
Für die Qualifizierung der OperateurInnen bedeutet die hohe Wahrscheinlichkeit von *complacency* und *automation Bias*, dass ein umfangreiches Training, das weit über eine theoretische Unterrichtung hinaus geht, notwendig ist. Öffentlicher Raum ist sehr viel komplexer als das beschriebene Lebenserhaltungssystem im Experiment. Da Trainings nicht nur einzelne Automationsfehler adressieren dürfen, stellt sich die Gestaltung und Durchführung eines auf die Vielseitigkeit des öffentlichen Raumes eingehenden Trainings als besonders zeitintensiv und kostspielig dar.

5.2.3 Zusammenfassung

Die Rolle und Verantwortung der OperateurInnen stellt sich bei Automatisierung als sehr ambivalent dar. Einerseits sollen Menschen im Überwachungsprozess weitestgehend ersetzt oder zumindest soweit assistiert werden, dass Kosten durch geringere Qualifikation oder Arbeitszeit reduziert werden können. Andererseits wird ihnen alle Verantwortung auferlegt, die ein System wie Ω nicht leisten kann, um ein hohes Maß an Automatisierung zu legitimieren. Diese Verantwortung wiederum verlangt ein Maß an Qualifikation, Reflexion und Mündigkeit, das mit dem ersten Ziel in Konflikt steht. Zusätzlich dazu werden und vor allem können den OperateurInnen nicht in dem Maße Informationen bereitgestellt werden, dass mündige Entscheidungen, besonnene Beurteilungen und die Überprüfung des Systems möglich wären. Selbst wenn OperateurInnen für eine konkrete Maßnahme nach bestem Wissen geschult werden, kommt es aus Mangel an Verifizierbarkeit und *complacency* zu Fehlentscheidun-

gen und somit zu Konsequenzen, die allein auf die automatisierte Verarbeitung des Ω-ähnlichen Systems zurückzuführen sind. OperateurInnen können nicht als Legitimation einer umfangreichen Automatisierung akzeptiert werden.

5.3 Diskriminierung der Betroffenen

Niemand darf wegen seines Geschlechtes, seiner Abstammung, seiner Rasse, seiner Sprache, seiner Heimat und Herkunft, seines Glaubens, seiner religiösen oder politischen Anschauungen benachteiligt oder bevorzugt werden. Niemand darf wegen seiner Behinderung benachteiligt werden.
ARTIKEL 3, ABSATZ 3 DES GRUNDGESETZES

Im Folgenden wird beschrieben, auf welche Weise diskriminierende Muster in ein System wie Ω gelangen können. Dies kann erwünscht sein und gezielt zur Risikoabschätzung eingesetzt werden, oder aber unbeabsichtigt geschehen. Im letzteren Fall besteht die Gefahr, dass die diskriminierende Überwachung aus den oben genannten Gründen weder von den OperateurInnen entdeckt und gemeldet wird, noch auf einfache Weise behoben werden kann, oder dass auf Grund der Komplexität der Strukturen eine Korrektur gar nicht möglich ist. Da auf diese Weise eine Diskriminierung wahrscheinlich wird, die automatisiert und daher institutionalisiert ist, ist mit noch umfangreicheren Rückkoppelungseffekten zu rechnen, als sie bei *manueller Videoüberwachung* auftreten konnten.

5.3.1 Politik von Technik

Meistens wird angenommen, Technik sei neutral und wertfrei.[285] Aufgrund dieser vermeintlichen Eigenschaft wird der Einsatz von Technik häufig als Lösung propagiert, um über menschliche Unzulänglichkeiten wie Voreingenommenheit und Diskriminierung hinweg zu kommen. Der Einsatz von *datahiding* wird beispielsweise als umfassende Reduzierung der Diskriminierung durch OperateurInnen beschrieben, was den Anschein der Neutralität des Gesamtsystems erweckt.[286] Computer werden

285 **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems", S. 195.
286 **Senior**, *Protecting Privacy in Video Surveillance*, S. 46.

im Zusammenhang mit automatisierter Videoüberwachung sogar explizit als indifferent z. B. bezüglich Alter und Geschlecht beschrieben.[287]
Technik *ist* jedoch politisch, denn mit ihrer Ausgestaltung werden bestimmte Interessen berücksichtigt und andere ausgeschlossen.[288] Mitunter ist diese „Mikropolitik" (*micro-politics*) der Technik nicht beabsichtigt (*authored*), sondern implizit und Teil eines oftmals profanen Prozesses bei der Lösung eines praktischen Problems.[289] Da Artefakte niemals isoliert wirken, sondern in einem *sozio-technischen Netzwerk* eingebettet sind, kann sich diese „Mikropolitik" auf unerwartete Weise multiplizieren und verstärken.[290]

5.3.2 Diskriminierung durch Algorithmen

Introna und Wood untersuchten diese „Mikropolitik" für Algorithmen beispielhaft an Gesichtserkennung.[291] Es wurde auf Grund der Funktionsweise der Algorithmen vermutet, dass jene Menschen besser erkannt werden, die stärker von „Normalität" abweichen und daher häufig zu Minderheiten gehören.
Tatsächlich konnten mehrere Studien bei mehr als zehn Gesichtserkennungsalgorithmen einen signifikanten *Bias* bezüglich Geschlecht, Alter und „Rasse" (*race*) nachweisen.[292] Asiaten sind leichter zu erkennen als Weiße, Afroamerikaner leichter als Weiße, ältere Menschen leichter als Junge und Menschen mit andersgearteter Haut (*other skin*) leichter als Menschen mit reiner Haut. Diese Unterschiede in den Erkennungswahr-

287 **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 26.
288 **Winner**, „Do Artifacts Have Politics?"
289 **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems", S. 197.
290 **Introna** ; **Nissenbaum**, „The Internet as a democratic medium: why the politics of search engines matters".
291 **Introna** ; **Wood**, „Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems".
292 Ebd., S. 190.

scheinlichkeiten liegen in einer signifikanten Größenordnung von 5% bis 10%. Introna und Wood diskutieren, welchen Effekt dies hat. Bei höheren Schwellwerten – z. B. als Maßnahme gegen zu viele Fehlalarme – würde die Wahrscheinlichkeit, dass diese Gruppen einen Alarm auslösen, höher sein als für andere Gruppen, wodurch sie auch bei einem Fehlalarm mit höherer Wahrscheinlichkeit eingehender untersucht werden als andere. Wie bereits in 2.3.3 beschrieben, kann sich diese Diskriminierung über Rückkoppelungseffekte über längere Zeiträume auch über die Grenzen des Systems hinweg manifestieren und intensivieren.
Derartige Probleme festzustellen, ist bei Algorithmen wie denen zur Gesichtserkennung besonders schwierig, da es sich um sogenannte „silent technology" handelt.[293] Stille Technik zeichnet sich aus durch ihre Unsichtbarkeit, Eingebettetheit, Flexibilität und Undurchsichtigkeit. Zusätzlich zur Tatsache, dass Algorithmen oft proprietär und daher nicht einsehbar sind, können sie auf Grund ihrer Komplexität auch nur von Experten verstanden werden.

5.3.3 Diskriminierung durch Ω

Ω ist mit dem Einsatz im öffentlichen Raum in einem besonders komplexen sozio-technischen Netzwerk eingebettet. Es steht durch Verarbeitung der *extrahierten* Informationen, *Projektion* in den Raum und der Reaktionen auf Auswertungsergebnisse in unmittelbarer intensiver Wechselwirkung mit dem überwachten Raum und kann indirekt auch längerfristige Auswirkungen auf die Gesellschaft haben. Für *manuelle Videoüberwachung* konnte nachgewiesen werden, dass es im Einsatz zu Diskriminierung kommt und diese sich über die Grenzen der Maßnahmen hinweg dauerhaft auswirken und manifestieren kann. Im Folgenden werden

293 Ebd., S. 183.

Überlegungen angestellt, ob es zu derartigen Effekten auch beim Einsatz von Ω kommen kann.

Präventiver Charakter und Diskriminierung

An automatisierte Videoüberwachung wird der Anspruch gestellt, nicht nur unerwünschtes Verhalten (z. B. Kriminalität und Gewalt) zu erkennen, sondern diese noch vor ihrem Auftreten zu prognostizieren. Um die *Sensitivität* zu erhöhen, wird *risk profiling* betrieben.[294] Zwei zentrale Charakteristika von *risk profiling* sind, dass es „hypothetische Information abgeleitet von angesammelten Daten" enthält und einen „proaktiven Charakter" hat.[295] Ersteres zeichnet sich dadurch aus, dass nicht nur individuelle Informationen ausgewertet werden, sondern statistische Informationen, abgeleitet von „riesigen"[296] Datenbanken, genutzt werden. Der proaktive Charakter ergibt sich daraus, dass nicht Spuren einer Straftat gefunden werden sollen, sondern dass Kriminalität bzw. unerwünschten Geschehnissen vorgebeugt werden soll. *Risk profiling* liefert dementsprechend keine Beweise sondern nur ein Ergebnis.

Die Vorgehensweise beim *risk profiling* steht im Konflikt mit dem Prinzip des Nicht-Diskriminierens (*non-discrimination*)[297], denn diese entspricht eben dem, was Mark Galliker und Franc Wagner als *soziale Diskriminierung* verstehen.[298] Dabei geht es um die rein kategorische Benachteiligung von Personen aufgrund einer (meist negativen) Beurteilung. Nach der Bewertung als Teil einer Kategorie werden die „Personen unter Absehung von ihren je besonderen Eigenschaften, Interessen und Verdiensten auf bloße Vertreter einer Kategorie reduziert".[299]

---

294 **Čas**, „The relevance of social and economic costs of surveillancy – Conclusion", S. 253.
295 **Fidis WP6 (Hrsg.):** *D6.7c: Forensic Profiling*.
296 Eigene Übersetzung.
297 **Čas**, „The relevance of social and economic costs of surveillancy – Conclusion", S. 253.
298 **Galliker ; Wagner**, „Ein Kategoriensystem zur Wahrnehmung und Kodierung sprachlicher Diskriminierung".
299 Ebd., S. 34.

Die Diskriminierung wird beim *risk profiling* nicht nur akzeptiert sondern bewusst als Mittel eingesetzt. Im Falle eines umfangreichen Einsatzes von Systemen wie Ω, die nach ähnlichen oder gleichen Kriterien kategorisieren, kommt es zu einer *Institutionalisierung* der Diskriminierung. Im Vergleich zur mehr oder weniger individuellen Diskriminierung durch OperateurInnen bei *manueller Videoüberwachung* ist daher mit einer signifikant stärkeren Manifestation und Verstärkung der Vorurteile, der Verzerrung von Statistiken und einem höheren Maß von Benachteiligung für Betroffene zu rechnen.

Wahrnehmbare und verdeckte Benachteiligung

Die Benachteiligung, die beim Einsatz von Ω für diskriminierte Personen zu erwarten ist, können für Betroffene wahrnehmbar sein, sich jedoch auch deren Kenntnis entziehen. Wahrnehmbare – aber nicht unbedingt für die Person als diskriminierend erkennbar oder empfundene – Nachteile, sind z. B. Maßnahmen, die vom System oder den OperateurInnen eingeleitet werden. Dies könnten Befragungen, Personenkontrolle, Durchsuchungen oder das Verweigern des Aufenthalts in einem bestimmten Bereich sein. Je nach Einsatzgebiet von Ω sind jedoch wesentlich drastischere Maßnahmen und Auswirkungen für Betroffene denkbar.

Von Betroffenen nicht ohne Weiteres wahrnehmbare Benachteiligung ist der tiefere Eingriff in die Persönlichkeitsrechte, wenn eine Person vom System oder den OperateurInnen einer eingehenderen Beobachtung unterzogen werden. Vergleichbar mit einer höheren Stufe des *Drei-Stufen-Modells*[300], werden mehr Daten über eine Person erhoben und intensiver ausgewertet. Es findet somit ein größerer Rechtseingriff statt. Dies ist auch dann als Benachteiligung zu bewerten, wenn der Person kein *direkter* Schaden entsteht. Eine zusätzliche Benachteiligung entsteht, wenn personenbeziehbare Informationen über den „Vorfall" gespeichert wer-

300 Siehe Kapitel 3.6.4.

den. Ein spürbarer Nachteil entsteht dann eventuell durch rückwärts gerichtete Überwachung.

5.3.4 Einfallstore für Diskriminierung

Während *risk profiling* ganz offensichtlich intendiert zu Diskriminierung führt, da sie bewusst als Mittel der sozialen Sortierung eingesetzt wird, gibt es auch weitere Einfallstore für Diskriminierung, die unter Umständen nicht beabsichtigt, oder zumindest nicht Teil der Gestaltung des *risk profiling* selbst sind. Die „Mikropolitik" einzelner Algorithmen und Module können ebenso dazu führen. Ein einfaches Beispiel ist die Annahme über menschliche Körper beim *region based tracking*. Der Körper eines Menschen wird erwartet als „eine Kombination einiger *blobs*, jeweils entsprechend verschiedener Körperteile wie Kopf, Torso und den vier Gliedmaßen."[301] Menschen, die diesem als normal angesehenen Körperbau nicht entsprechen, etwa weil ihnen eines der Gliedmaßen fehlt, würden je nach dem, wie mit solchen Abweichungen umgegangen wird, bestimmte weitere Prozesse anstoßen. Ein denkbares Resultat ist das Alarmieren der OperateurInnen, die die „Situation" einschätzen sollen und den Alarm entsprechend ignorieren. Fließt das abweichende Ergebnis jedoch nur in eine Kette weiterer z. B. stochastischer Berechnungen mit ein, ist denkbar, dass die Person den OperateurInnen mit einem höheren Level von Verdächtigkeit oder anderen Eigenschaften präsentiert wird, die von ihnen nicht mehr unmittelbar mit der wahrnehmbaren Tatsache des fehlenden Körperteils in Verbindung gebracht werden können, da sie wie oben dargestellt nicht über die genaue Herleitung der Ergebnisse informiert sind. Eine andere Möglichkeit der Diskriminierung für die Betroffenen ist wiederum die Tatsache gesteigerter Datenerhebung. Kann etwas

301 Eigene Übersetzung mit eigens vorgenommener Hervorhebung, **Hu** et al., „A survey on visual surveillance of object motion and behaviors", S. 338.

von Ω nicht genau bewertet werden, so werden mehr Daten erhoben, um die Einschätzung zu verbessern. In einer überwachten Welt würden derartig „auffällige" Menschen andauernd einer gesteigerten Datenerhebung und eben auch einer intensiveren Kontrolle und Bewertung ausgesetzt als Menschen, die den Annahmen entsprechen.
Neben solchen zu Diskriminierung führenden Annahmen beim Entwurf von Algorithmen, liegen Einfallstore vor allem bei den Modellen, über deren Übereinstimmung (*Blacklistansatz*) und Abweichung (*Whitelistansatz*) mit dem beobachteten Verhalten, Auffälligkeit erkannt werden soll. Der *Whitelistansatz* führt generell zu Diskriminierung. Durch Modellierung von „Normalität" werden jene Menschen einer eingehenderen Untersuchungen unterzogen die von diesem Modell abweichen. Die Wahrscheinlichkeit, dass jene Menschen zu einer Minderheit gehören und auch ohne den Einsatz von Systemen wie Ω Diskriminierung ausgesetzt sind, ist dabei gesteigert.
Modelle für beide Ansätze können wie in 3.1.3 beschrieben einerseits manuell, andererseits automatisiert über beaufsichtigtes und unbeaufsichtigtes Lernen erzeugt werden. Alle drei Varianten stellen Einfallstore für Diskriminierung dar.

Manuelle Erzeugung der Modelle

Bei der manuellen Erzeugung von Modellen besteht die Gefahr, dass jene Kriterien genutzt werden, die auch schon bei *manueller Videoüberwachung* zu Diskriminierung führten.
Es handelt sich dann nicht mehr um (z. B. rassistische) Vorurteile der OperateurInnen, sondern um die Vorurteile und Annahmen der EntwicklerInnen, sowie der BetreiberInnen und BenutzerInnen, die die Mo-

delle für ihren Einsatz selbst wählen oder erstellen können.[302] Auch älterer Code oder Code-Bibliotheken können schon einen Bias enthalten.[303] Wie beschrieben, geht der Trend jedoch dahin, die Erstellung der Modelle und Kriterien zu automatisieren.

Modellgewinnung mit beaufsichtigtem Lernen

Beim beaufsichtigten Lernen bestehen Einfallstore neben den Annahmen und Entscheidungen, die bei der Implementierung des Lernverfahrens getroffen wurden, besonders in der Auswahl der Trainingsdaten, die aus einer Menge von Eingaben und den erwarteten Ausgabewerten bestehen. Enthalten diese Trainingsdaten beabsichtigt oder unbeabsichtigt Verzerrungen der Häufigkeiten bestimmter Eigenschaften, so ist zu erwarten, dass auch die Modelle diese Verzerrungen beinhalten. Deutlich wird der Effekt, wenn man den Extremfall betrachtet, bei dem nur ein extrem begrenztes Spektrum einer ansonsten facettenreichen Eigenschaft in den Trainingsdaten vorkommt. Soll ein Modell beispielsweise einen „normalen" Laufstil oder „normale" Gestik repräsentieren (*Whitelistansatz*) und werden Modelle nur anhand der Bewegung europäischer Menschen trainiert, so würden Personen aus anderen Kulturen anhand ihrer Bewegungen als alarmierend und deviant bewertet werden. Auch ein Modell für den *Blacklistansatz* könnte verzerrt sein.[304] Soll eine aggressive Körpersprache erkannt werden und würde diese ausschließlich anhand des Verhaltens einer bestimmten Auswahl von Menschen modelliert werden, kann ähnliche, aber ungefährliche Körpersprache anderer Menschengruppen als aggressiv bewertet werden. Anhand eigener Erfahrung kann man feststellen, dass auch innerhalb einer Stadt wie z. B. Berlin so unterschiedliche Menschen mit einer breiten Palette von Ver-

302 Vgl. **Musik**, „The thinking eye is only half the story: high-level semantic video surveillance", S. 348.
303 Vgl. **Macnish**, „Unblinking eyes : the ethics of automating surveillance", S. 15.
304 Vgl. **Kleinz**, *Generalverdacht der Algorithmen*.

haltensmustern leben, die auf unterschiedlichste Weise zu interpretieren sind, dass kein Trainingsdatensatz adäquat sein kann. Darüber hinaus ist das Erlangen von Trainingsdatensätzen problematisch.[305] Für bestimmte, eigentlich besonders im Fokus stehende Szenarien wie Terrorismus oder das Auskundungsverhalten von Bankräubern, existieren keine umfangreichen Trainingsdaten.[306] Teilweise werden daher eigens SchauspielerInnen engagiert, die den EntwicklerInnen das benötigte Material erzeugen. Bei Versuchen der Mimikerkennung wurden die Trainingsdaten nicht nur von Experten für Mimik sondern größtenteils auch von Laien erstellt und ausgewählt.[307] Dass derartig gewonnene Trainingsdaten frei von Vorurteilen und Annahmen ist, ist zu bezweifeln. Man könnte hier argumentieren, dass dieser Bias in der Entwicklungsphase nicht zum Tragen kommen würde, da die Modelle nur für den Test verwendet werden würden. Tatsächlich werden jedoch auch die Lernalgorithmen an die Trainingsdaten angepasst. So ist beispielsweise bei neuronalen Netzen die Topologie und die genaue Ausprägung des Lernverfahrens von entscheidender Bedeutung für das spätere Lernergebnis. Finden derartige Algorithmen als proprietäre Software in einem System wie Ω Anwendung, so werden Informationen über derartige Voraussetzungen und Annahmen wahrscheinlich nicht mitgeliefert. Die Möglichkeit des eigenständigen Antrainierens der Software durch die BenutzerInnen täuscht dann über einen eventuellen Bias der Lernverfahren oder der Strukturen hinweg.

305 **Dee** ; **Velastin**, „How close are we to solving the problem of automated visual surveillance?“, S. 339.

306 **Musik**, „The thinking eye is only half the story: high-level semantic video surveillance“, S. 348.

307 Ebd., S. 347.

Modellerzeugung mit unbeaufsichtigtem Lernen

Bei unbeaufsichtigtem Lernen, besonders wenn dieses zur Laufzeit des Systems stattfindet, besteht eine gesteigerte Gefahr für die Entstehung diskriminierender Modelle. Da versucht wird, Muster in Eingabedaten zu erkennen, ohne das Vorgaben gemacht werden, können für Verdächtigkeit und Normalität Kriterien aus den Daten gewonnen werden, die sich der Beobachtbarkeit, der Nachvollziehbarkeit oder zumindest der Aufmerksamkeit eines Menschen entziehen. Selbst bei gewissenhaft ausgewählten Trainingsdaten könnten sich auf diese Weise diskriminierende Kriterien herausbilden. Diese Kriterien müssen jedoch längst nicht so „eindrucksvoll", alarmierend und eindeutig sein, wie etwa die Hautfarbe. Die Grenzen, die hier automatisiert gezogen werden, müssen nicht entlang von Kulturkreisen oder Ethnien verlaufen.

Nichtvisuelle Kriterien für Diskriminierung

Ohne explizit darauf einzugehen, konzentrierten sich die letzten Ausführungen fast nur auf visuell beobachtbares Verhalten. In 3.3.1 und3.4.1 wurde jedoch bereits beschrieben, dass von Ω auch andere Informationsquellen wie Mikrofonsignale, systeminterne Datenbanken und externe Daten genutzt werden können. Dies soll an dieser Stelle erwähnt werden, um den Blick zu weiten, welche weiteren Kriterien potentiell zu Diskriminierung führen können. Hat Ω Zugriff auf Biometriedatenbanken, Melderegister, Strafregister und Internetplattformen wie soziale Netzwerke, Blogs oder Kommunikationsdienste, so sind dem Umfang der Kriterien für Diskriminierung – zumindest technisch – kaum noch Grenzen gesetzt. Führt beispielsweise die Mitgliedschaft in einem Selbstverteidigungsverein in einem Bezirk mit erhöhter Kriminalität in Kombination mit einem vermeintlich aggressiven Schreibstil auf Twitter dazu, dass den OperateurInnen für den betreffenden Menschen ein höheres Risikopotential angezeigt wird, so findet auch hier eine Diskriminierung statt.

Die Person wird der Gruppe der „potentiell aggressiven und kriminellen KampfsportlerInnen“ zugeordnet und erfährt dadurch den Nachteil, eingehender beobachtet oder anderer Maßnahmen unterzogen zu sein, die sich ausschließlich aus der Kategorisierung ergeben.

5.3.5 Diskussion von Gegenmaßnahmen

Kevin Macnish zieht in Erwägung, dass automationsbedingte Diskriminierung im Gegensatz zu Diskriminierung durch OperateurInnen, ein geringeres Problem darstellen könnte.[308] Dafür werden zwei mögliche Gründe angeführt. Erstens könnten die OperateurInnen, die die Diskriminierung bemerken, entweder diese Tatsache melden oder die Hinweise des Systems, die auf dem Vorurteil basieren, ignorieren. Zweitens seien Vorurteile womöglich „leichter zu korrigieren“ als Vorurteile von einem oder mehreren OperateurInnen. In einer „simplen roll-out Prozedur“ könnten die Vorurteile „über Nacht“ im gesamten System „ausgemerzt“ („eradicate“) werden.[309] Beide Argumente gehen davon aus, dass die Diskriminierung sich 1) scharf und eindeutig darstellt und auch technisch abgrenzbar ist und 2) von OperateurInnen registriert werden.
Eine *Eindeutigkeit* möglicher Diskriminierungskriterien ist unwahrscheinlich. Dass automatisierte Diskriminierung anhand von Kriterien entstehen kann, die sich einem Mensch nicht unmittelbar erschließen, wurde bereits diskutiert. Auch die eindeutige „Lokalisierung“ der Ursachen im System ist ein schwieriges, vermutlich sogar unlösbares Problem. Zwar wurde bisher die Formulierung „Bias eines Modells“ oder „eines Moduls“ benutzt, was eine gewisse Beschreibbarkeit und Abgeschlossenheit impliziert, doch wie in 3.1.1 bereits beschrieben, liegen die Modelle mitunter nicht als klar abgegrenzte Datensätze vor, sondern ergeben sich mitunter

308 **Macnish**, „Unblinking eyes : the ethics of automating surveillance“.
309 Ebd., S. 18.

aus Zusammenhängen, die über mehrere Abstraktionslevel verteilt sein können. Modelle können außerdem auf Basis der statistischen Auswertung unzähliger Einzelinformationen gebildet worden sein und untrennbar miteinander verknüpft (etwa in einem einzigen neuronalen Netz) vorliegen. Es ist davon auszugehen, dass weder die Bedeutung noch die genaue Entstehung einzelner Parameter beschrieben werden können. Die einfache „Ausmerzung" und Korrektur einzelner diskriminierender Modelle oder gar einzelner Aspekte eines Modells ist in der Praxis kaum umzusetzen.

Auch die *Registrierbarkeit* – der zweite Ausgangspunkt von Macnish für einfache Korrigierbarkeit – ist strittig.[310] OperateurInnen müssten dazu 1) wissen, auf welche Weise und nach welchen Kriterien der Alarm oder die angezeigte Information entstanden ist, 2) in der Lage sein diese Informationen zu deuten und 3) verantwortungs- und selbstbewusst genug sein, die nötigen Konsequenzen daraus zu ziehen. In 5.2.1 wurde bereits argumentiert, dass die Erfüllung dieser Bedingungen von den OperateurInnen nicht erwartet werden kann.

Diskriminierung äußert sich nicht in einem einzelnen Fall, so dass OperateurInnen auf ein mal „stutzig" werden. Es ist eher zu erwarten, dass Tatsachen wie das häufige Warnen z. B. vor älteren Asiaten mit vernarbtem Gesicht (vgl. Introna und Wood) nicht als Diskriminierung wahrgenommen wird, sondern dazu führt, dass dieses Vorurteil sich bewusst oder unbewusst auch in den OperateurInnen manifestiert, die tagtäglich mit dem System arbeiten. Durch eine mögliche, durch Verzerrung erhöht wahrgenommene Kriminalität der Betroffenen, werden die OperateurInnen in ihrem Vorurteil bekräftigt. Wer nicht von der konstruierten „Normalität" abweicht, wird nicht vom System markiert und fällt somit aus dem Fokus der OperateurInnen.

310 **Macnish**, „Unblinking eyes : the ethics of automating surveillance".

Auch dieses spricht ergänzend zu den Überlegungen in 5.2 eher für die Notwendigkeit gut ausgebildeter OperateurInnen, als für die Möglichkeit durch Automatisierung die menschliche Verantwortung an ein System wie Ω abzutreten. Aus ähnlichen Gründen zieht auch Christoph Musik diesen Schluss.[311] Dort wird auf Grund der zu starken Vereinfachung und der damit verbundenen geringen Effektivität der Bedarf festgestellt, dass OperateurInnen sowohl Verfahren als auch die Trainingsdaten („ground truth") kennen müssen, um mit der Technik umgehen und Risiken minimieren zu können.[312]

311 **Musik**, „The thinking eye is only half the story: high-level semantic video surveillance".

312 Ebd., S. 351.

5.4 Wirkung der Automatisierung auf Individuen und Gesellschaft

Niemand interessiert, wofür etwas „gedacht" ist. Die relevante Frage ist, wer davon betroffen ist.[313]

FELIX VON LEITNER

5.4.1 Informationsasymmetrie

Im Kapitel 2.3.3 wurde *manuelle Videoüberwachung* in einen Zusammenhang mit Foucaults Disziplinierungsthese nach dem panoptischem Prinzip gebracht. Effekte sind vor allem eine selbstdisziplinierende Wirkung auf Individuen und eine steigende Konformität der Gesellschaft. Ursache dieser Effekte ist die Informationsasymmetrie der Kontrolle – gesehen zu werden, ohne überprüfen zu können, ob und in wieweit eine Überwachung stattfindet.

Dass dieses Ungleichgewicht durch den Einsatz von Systemen wie Ω gegenüber *manueller Videoüberwachung* beträchtlich verstärkt wird, wurde in Kapitel 5.1 hergeleitet. Das Wissen über die Überwachten ist wesentlich umfangreicher, während die Informiertheit der Beobachteten über die genaue Ausgestaltung der qualitativ und quantitativ enorm erweiterten Möglichkeiten gering bleibt. Da darüber hinaus der Umfang der Überwachung immer weiter zunimmt, ist zu erwarten, dass auch die selbstdisziplinierende Wirkung und der Effekt der Konformität sich beim Einsatz automatisierter Videoüberwachung verstärken.

Wenn „allein schon" „die anlasslose Speicherung von Telekommunikationsverkehrsdaten geeignet [ist], ein diffus bedrohliches Gefühl des Beobachtetseins hervorzurufen, das eine unbefangene Wahrnehmung der

313 Quelle: URL:`http://blog.fefe.de/?ts=af41e7b1` (29.06.2013).

Grundrechte in vielen Bereichen beeinträchtigen kann“[314] und daher in der angestrebten Form vom Bundesverfassungsgericht für grundgesetzwidrig erklärt wurde, so ist auch die Wahrnehmung der Grundrechte beim Einsatz von Systemen wie Ω oder auch nur Teilen von Ω gefährdet.

5.4.2 Unterschied zwischen *manueller* und automatisierter Videoüberwachung

Von Interesse ist, welche Unterschiede es in der Auswirkung auf Betroffene macht, wenn diese wissen, dass statt eines Menschen ein „Computer“ die Bilder auswertet. Einem Aspekt dieser Fragestellung gingen Psychologen des Forschungsprojektes *MuViT*[315] nach. Sie überprüften empirisch die sogenannte „Einschüchterungsthese“, indem getestet wurde, ob sich das Verhalten von Versuchspersonen ändert, wenn sie glauben, gerade von mustererkennenden Kameras beobachtet zu werden. Ohne, dass bisher genauere Ergebnisse öffentlich vorliegen, wurde geäußert, dass viele Versuchspersonen in diesen Laborstudien ihr Verhalten verändert haben, eine gesteigerte Selbstwahrnehmung hatten und die Überwachung, wenn möglich, vermieden.[316]

Unterschied der Auswertungsquantität

Während das Wissen über „Mustererkennung“ auf Grund von Unkenntnis über die genauen Umstände ein recht diffuses Gefühl der Unsicherheit erzeugt, können Szenarien für bestimmte Techniken, die in Ω enthalten sind, ganz konkret entworfen werden. Während man sich beispielsweise bei *manueller Videoüberwachung* in der Masse der Bilder begründe-

314 BVerfG, 1 BvR 256/08 vom 2.3.2010, Absatz-Nr. (1 - 345), `http://www.bverfg.de/entscheidungen/rs20100302_1bvr025608.html`.

315 *MuViT* steht für Mustererkennung und Video Tracking: sozialpsychologische, soziologische, ethische und rechtswissenschaftliche Analysen.

316 **Becker**, *Automatische Absichtserkennung*.

ter Weise in einer gewissen Anonymität wähnen konnte, ist diese mit der umfangreichen, auch möglichen zeitlich rückwärts gerichteten Auswertung in Kombination mit Biometrie, nicht mehr gegeben. Folgen des *chilling effects*, z. B. nicht mehr an Ereignissen wie Demonstrationen teilzunehmen, wenn nicht ausgeschlossen werden kann, dass die Teilnahme zu Nachteilen führt oder später zu Nachteilen führen könnte, kann durch die automatisierte Identifizierung und nachträgliche Auswertung verstärkt oder überhaupt erst erzeugt werden.

Google hielt es auf eine Anfrage der US-Regierung hin für technisch möglich, alle privaten Kameras so zu vernetzen, dass staatliche Behörden jederzeit auf sie zugreifen könnten.[317] Werden derartige Schnittstellen etwa für Geheimdienste zur Verfügung gestellt, gewinnt die Theorie des *chilling effects* ein großes Stück an Bedeutung und Realität. In Anbetracht der Überwachungsprogramme PRISM und *Tempora*, deren Ziel es ist, gewisse Daten für immer vorrätig zu halten, sind genau die dystopischen Szenarien, die mit der Theorie des *chilling effect* entworfen werden, technisch ermöglicht. Anhand der Dimensionen des Datenzentrums der NSA in Utah mit nahezu 9.300 Quadratmetern Fläche für Hochleistungsrechner[318] und einer geschätzten Speicherkapazität von ca. fünf Milliarden Terabyte[319] erscheint die Verarbeitung derartiger Datenmengen denkbar.

Auswirkungen trotz eingeschränkter Fähigkeiten des Systems

Da der Großteil einer Bevölkerung nicht in der Lage ist, technische Möglichkeiten der Verhaltenserkennung und Datenauswertung angemessen einzuschätzen, besteht die Möglichkeit, dass die Systeme in ihrer Leistung auch von den Betroffenen überschätzt werden. Der Gebrauch von Begriffen wie *intelligente Videoüberwachung*, die in Kapitel 3 diskutiert wurden, wecken bestimmte Assoziationen und tragen zusätzlich zu un-

317 **Dix**, „Datenschutz und Informationsfreiheit : Bericht 2010", S. 10.
318 **Carroll**, *Welcome to Utah, the NSA's desert home for eavesdropping on America*.
319 **Beaupoil**, *Das Netzwerk der NSA : Horchposten – auch in Deutschland*.

angemessenen Vorstellungen bei. Ähnlich wie zum Ausbau *manueller Videoüberwachung* ist damit zu rechnen, dass Wirtschaft und Politik die Fähigkeiten und vermeintliche Effizienz der Systeme anpreisen werden – vermutlich sogar mit Erfolgsversprechen, die von verheißungsvollen Experimenten unter Laborbedingungen abgeleitet werden. Auch durch solche öffentlichkeitswirksamen Informationen entstehen unangemessene Vorstellungen der Leistungsfähigkeit. Videoüberwachung kann demnach auch dann eine gesteigerte disziplinierende Wirkung zeigen und die oben beschriebenen Effekte nach sich ziehen, wenn das von Betroffenen zu vermeiden angenommene Verhalten gar nicht Ziel der Überwachung ist, oder technisch (noch) gar nicht erkannt werden kann. Sogar bei manuellen Maßnahmen kann allein die Tatsache, dass Systeme wie Ω überhaupt zum Einsatz kommen, dazu führen, dass man nicht ausschließen kann, automatisiert überwacht zu werden, wenn sicheres Wissen über die genauen Umstände einer Maßnahme fehlt. Durch die Möglichkeit der Speicherung von Rohdaten, deren Weitergabe und Verarbeitung an anderer Stelle und zu einem späteren (eventuell sogar wesentlich späteren) Zeitpunkt, kann auch bei eigentlich *manueller Videoüberwachung* im Nachhinein noch eine rückwärts gerichtete automatisierte Überwachung durchgeführt werden.

5.4.3 Diskussion von Gegenmaßnahmen

Um die Effekte zu verringern, müssten Betroffenen Informationen über die genaue Ausgestaltung einer Maßnahme zur Verfügung gestellt werden. Während dies bei *manueller Videoüberwachung* Informationen über Betreiber, Umfang, Speicherungsdauer und Gründe der Maßnahme waren, müssten bei automatisierter Videoüberwachung weitaus mehr Informationen gegeben werden – etwa ob Gesichtserkennung oder Bewegungserkennung genutzt wird, nach welchen Modellen, Profilen und Kriterien

bewertet wird, ob und wie lange Rohdaten, Analyseergebnisse und Metadaten gespeichert werden usw. Außerdem müssten effektive, technische Maßnahmen zur Absicherung von Datenschutz, gegen einen potentiellen Missbrauch und gegen Leaking und Tapping zum Einsatz kommen. Maßnahmen wie *datahiding*, die keine Anonymisierung bewirken, sondern lediglich vor den Blicken der OperateurInnen schützen und nicht absolut die automatisierte Weiterverarbeitung verhindern, können hier kaum oder gar nicht helfen. Für die nötige Glaubhaftigkeit der Absicherungen müssten diese Maßnahmen außerdem nachprüfbar sein. Dazu müsste nicht nur die Software quelloffen zur Verfügung stehen, sondern auch die gelernten Modelle, die die Arbeitsweise im Wesentlichen beeinflussen, zugänglich und nachvollziehbar sein. Dass diese Voraussetzungen erfüllt werden, ist höchst unwahrscheinlich. Nicht nur, dass Code und Modelle schwer nachzuvollziehen wären – schon die Tatsache, dass derartige Software quelloffen zur Verfügung steht, ist aus wirtschaftlichen und aus Gründen der Sicherheit nach dem Prinzip *security by obscurity* nicht zu erwarten.

5.4.4 Quantitätsproblem wird zu Qualitätsproblem

Die Probleme sind also nicht für einzelne Kameras oder einzelne Überwachungsmaßnahmen zu lösen. Die beschriebenen Effekte entstehen allein schon durch die Tatsache, dass umfangreiche manuelle und prinzipiell auch automatisierte Videoüberwachung stattfindet, als gesamtgesellschaftlich wirksames Phänomen und daher gesamtgesellschaftliches Problem. Selbst wenn jede einzelne Maßnahme in ihrer Ausprägung vollkommen mit dem Recht vereinbar wäre[320], erzeugen die Kameras den Eindruck einer ortsübergreifenden oder in Städten sogar flächendecken-

320 Es sei an die in Kapitel 2.1 erwähnten, alarmierenden Zahlen über die Rechtskonformität öffentlicher Videoüberwachungsanlagen erinnert.

den Überwachung und erzeugen eine Stimmung des Misstrauens und des Verdachtes. Mit dem Befürchten (oder im dystopischsten Fall dem Wissen) von kamera- und sogar systemübergreifender Verfolgbarkeit von Personen und der fehlenden Überprüfbarkeit, ob und wie lange diese Daten vorgehalten werden, muss theoretisch davon ausgegangen werden, dass jede Regelübertretung und jedes unerwünschte Verhalten gespeichert wird. Mit der Möglichkeit beliebig weit rückwärts gerichteter Überwachung können Kriterien der Regeln für Unerwünschtheit sogar im Nachhinein noch festgelegt werden.
Während also den Individuen immer weniger vertraut wird – Überwachung ist als Zeichen des Misstrauens zu verstehen – müssen diese den staatlichen und privaten Betreibern der Maßnahmen immer mehr Vertrauen entgegenbringen. Will man am gesellschaftlichen Leben im öffentlichen Raum teilnehmen, kann man sich – wie z. B. in London – der Überwachung schon jetzt nicht mehr entziehen.

5.4.5 Einsatz außerhalb eines demokratischen Rahmens

Zwar lässt die Überwachungspraxis der NSA und des britischen Geheimdienstes, die Internetverkehr und somit auch die Kommunikation von Jedem und Jeder überwachen, und die verhaltenen Reaktionen der Bundesregierung den Schutz der Grundrechte in Frage stellen, ist wohl doch auf den grundsätzlichen Schutz der Menschen- und Grundrechte in Europa zu hoffen. Ω wurde mit dem vollen technischen Potential entworfen, das viele Elemente enthält, die mit deutschem oder europäischen Recht nicht vereinbar ist. Dies bedeutet jedoch nicht, dass eine eingeschränktere Version von Ω die identifizierten Probleme nicht auch verursachen könnte. Die Probleme ergeben sich hauptsächlich aus der grundsätzlichen mit dem heutigen Stand der Technik gar nicht wesentlich anders zu gestaltenden Funktionsweise von Ω.

In anderen Ländern sind die Menschen- und Grundrechte jedoch wesentlich unzuverlässiger geschützt oder sind praktisch nicht vorhanden. In solchen Ländern ist der Einsatz von Ω in der oben beschriebenen Form durch Machthabende ein realistisches Szenario. Ohne rechtlich eingeschränkt zu sein, kann Ω bzw. die Vernetzung vieler Instanzen von Ω als ein gefährliches und menschenrechtsverachtendes Instrument der Macht genutzt werden. Wird, wie bei bei INDECT, trotz der Unvereinbarkeit mit dem Recht an Systemen wie Ω geforscht, besteht die Gefahr und die ökonomische Begehrlichkeit, derartige Überwachungstechniken in eben diese Länder zu exportieren. Dafür spricht auch das Ziel des Bundesministeriums für Bildung und Forschung (BMBF), dass „innovative deutsche Unternehmen sowie forschende Einrichtungen von [dem] boomenden Markt [für Sicherheitsprodukte und -dienstleistungen] profitieren" sollen.[321]

Vergleichbares ist bereits geschehen: *Nokia Siemens Networks* lieferte Technik zur Überwachung von Mobilfunknetzen in den Iran. Laut Frontal 21 wurden diese Techniken eingesetzt, um Menschen Droh-SMS zu schicken, für die anhand der Bewegungsmuster der Telefone festgestellt wurde, dass sie sich auf dem Weg zu Protesten befanden.[322] Dies ist in Anbetracht der Möglichkeiten von Ω ein vergleichsweise geringes Maß von Überwachung.

Besonders hervorzuheben ist in diesem Zusammenhang, dass OperateurInnen keineswegs essentieller Bestandteil der Technik bzw. des Konzeptes sind. Für einen missbräuchlichen Einsatz kann die Bestätigung des Alarms durch einen Menschen ausgelassen werden und die weiteren Schritte direkt eingeleitet werden. Daniel Suarez stelle beispielsweise Selbstschussanlagen (*automated sniperstations*) vor, die Personen mit den

321 **BMBF (Hrsg.):** *Forschung für die zivile Sicherheit 2012 – 2017 : Rahmenprogramm der Bundesregierung*, S. 12.

322 **Frontal 21**, *Sendung vom 26.01.2010.*

DoDaam Super aEgis II

SGR A1 Samsung Techwin

Abb. 26: Automatisierte Schießanlage mit Techniken zur Bewegungserkennung, die auch in Ω zur Anwendung kommen.

gleichen Techniken erkennen und bewerten wie sie auch in Ω genutzt werden (Abb. 26).[323]

Systeme wie Ω zu exportieren, wenn ein derartiger Einsatz nicht auszuschließen ist, ist nicht nur menschenrechtsverachtend, sondern verhindert auch das Entstehen von Widerstand und Demokratie und gefährdet die Freiheit. Der Export und die konkrete Forschung an Technik, deren Einsatz unweigerlich gegen Grundrechte verstoßen würde, sollte stark reglementiert oder verboten werden. Momentan jedoch werden Forschungsprojekte wie etwa INDECT, deren Resultate mit europäischen Recht nicht vereinbar sind, mit Millionen von Euro von der Europäischen Union gefördert.

323 **Suarez**, *The kill decision shouldn't belong to a robot*.

6 Schluss

6.1 Zusammenfassung

Im Vorangegangenen wurden gesellschaftliche Probleme identifiziert, mit denen aus Perspektive der Informatik durch Automatisierung von Videoüberwachung gerechnet werden müssen. Außerdem wurden Lösungsansätze für die Probleme diskutiert.
Vorbereitend wurden relevante Aspekte und gesellschaftliche Probleme des Konzepts *manueller Videoüberwachung* dargelegt und Techniken zur Automatisierung beschrieben.
Videoüberwachung kann öffentlichen Raum in seiner Komplexität nicht erfassen. Die OperateurInnen überwachen daher eine Abstraktion des überwachten Raumes. Bei der Auswahl der näher zu betrachtenden Bilder, deren Interpretation und der Entscheidung über Konsequenzen, müssen sie die Informationen rekontextualisieren. Dabei spielen Erfahrung der OperateurInnen und ihr Wissen über den überwachten Raum eine entscheidende Rolle.
Die bei Videoüberwachung erhobenen Daten sind durch ihre Personenbeziehbarkeit datenschutzrechtlich relevant. Das Recht auf informationelle Selbstbestimmung wird jedoch durch eine starke Informationsasymmetrie zwischen Betroffenen und Überwachenden stark eingeschränkt.
Aus diesen Eigenschaften ergeben sich neben dem Verlust von *privacy* eine Reihe weiterer negativer Auswirkungen für Individuen und Gesellschaft. Videoüberwachung kann zu Kriminalisierung und durch Vorurteile der OperateurInnen zu Diskriminierung führen. Die starke Informationsasymmetrie führt zu einer Selbstdisziplinierung von Betroffenen und langfristig zu stärkerer Konformität der Gesellschaft. Videoüberwachung birgt damit eine Gefährdung gesellschaftlich-kultureller Entwicklung.

Ziel der *Automatisierung* von Videoüberwachung ist es, die Aufgaben der OperateurInnen zu unterstützen und in letzter Hinsicht möglichst vollständig zu automatisieren. Darüber hinaus sollen die Möglichkeiten der Computerisierung ausgenutzt werden, die eine qualitativ, aber vor allem quantitativ gesteigerte Erhebung, Verarbeitung und Nutzung der Daten ermöglicht, die von Menschen nicht geleistet werden kann.
Es wurden zahlreiche Techniken und Konzepte geschaffen, die für die Automatisierung verwendet werden sollen oder Teil eines automatisierten Videoüberwachungssystems sein können. Da es nicht genügt, einzelne Techniken auf Probleme hin zu untersuchen, sondern sich der Charakter automatisierter Videoüberwachung erst durch ihr Zusammenspiel erkennen lässt, wurde ein wahrscheinliches Komplettsystem Ω und dessen Funktionsweise anhand zugänglicher Informationen entworfen.
Ω besteht grundsätzlich aus einem Netzwerk von Kameras und Sensoren, Schnittstellen zu externen Informationsquellen, einem Analysemodul, Speichermöglichkeiten und einem Interface für die Mensch-System-Interaktion.
Die Sensoren übermitteln Daten an die Kameras, die dort mit den Bilddaten semantisch vorverarbeitet und an das Analysemodul übermittelt werden. Dieses integriert die Daten und wertet sie weiter aus. Personen können anhand von Gang und Gesicht identifiziert, über mehrere Kameras hinweg verfolgt und ihr Verhalten, Körpersprache und Mimik ausgewertet werden. Zusätzlich zu eigens erhobenen Daten können die externen Datenquellen wie Datenbanken oder Internet für weitere Informationen über Personen herangezogen werden. Analyseergebnisse und Rohdaten können zur späteren Nutzung und Verarbeitung gespeichert werden. Die Speicherung ermöglicht eine zeitlich rückwärts gerichtete Überwachung nach Kriterien, die im Nachhinein festgelegt werden können. Über eine externe Schnittstelle zu Ω kann eine solche auch von außen definiert und

angestoßen werden. Das Analysemodul stellt den OperateurInnen Bildmaterial und Analyseergebnisse sowie implizite und explizite Handlungsempfehlungen integriert in einer virtuellen interaktiven 3D-Umgebung dar. Ω übernimmt auf diese Weise eine auswählende, interpretierende und schlussfolgernde Funktion. Die Assistenz durch das System reicht damit weit in alle Aufgabenbereiche und Handlungsphasen der OperateurInnen hinein.

Zentrale Technik von Ω ist das Bildverstehen. Neben Objekterkennung und Techniken zur Identifizierung und Erfassung von Merkmalen von Personen, steht im Mittelpunkt der Überwachung die Verhaltenserkennung. Sie soll nicht nur eingesetzt werden, um unerwünschtes Verhalten zu detektieren sondern auch, um Prognosen über in der Zukunft liegende unerwünschte Geschehnisse zu ermöglichen. Erfassung und Interpretation von Verhalten und Geschehnissen in Bildmaterial wird mit einem komplexen Zusammenspiel verschiedenster Techniken vor allem aus Bildverarbeitung und künstlicher Intelligenz umgesetzt. Auf verschiedenen Abstraktionsebenen werden Bewegungen abgeschätzt und mit Modellen nachgebildet. Auf jeder Stufe werden Informationen auf Grund der Technik und der getroffenen Annahmen reduziert und vereinfacht. Die so erhaltenen Repräsentationen von Bewegungen werden dann mit vorher definierten Modellen verglichen, um interpretiert zu werden. Für die Erkennung von Unerwünschtem kann entweder die *Übereinstimmung* mit einem konkreten Modell von Unerwünschtem (*Blacklistansatz*) oder die *Abweichung* von einem Modell von „Normalität" (*Whitelistansatz*) gemessen werden. Ergebnis des Abgleichs, der unter Berücksichtigung von Abweichungen stattfinden muss, sind im Wesentlichen Wahrscheinlichkeitswerte, die auf höheren Abstraktionsebenen weiter verarbeitet und zu einer Interpretation genutzt werden. Ω entscheidet anhand der Inter-

pretation ob und auf welche Art und Weise die generierte Information den OperateurInnen angezeigt wird.
Ein weiterer relevanter Aspekt ist die Art der Gewinnung der Modelle und die Form, in der sie anschließend vorliegen. Da eine manuelle Erstellung aufwendig ist, werden eher Verfahren des Maschinenlernens zur automatischen Erzeugung genutzt. Aus Trainingsdaten ermitteln die Lernverfahren selbstständig Kriterien, nach denen die Daten beim Einsatz bewertet oder kategorisiert werden. Die resultierenden Kriterien und Modelle liegen mitunter in einer Form vor, die kaum oder gar nicht von einem Menschen verstanden oder manuell angepasst werden kann.
Aus diesen technischen und konzeptuellen Eigenschaften, den Zielen von Videoüberwachung und den Eigenschaften und Problemen *manueller Videoüberwachung* konnten Probleme abgeleitet werden, die sich aus der Automatisierung ergeben und Auswirkungen auf Individuen und Gesellschaft haben:
Automatisierung von Videoüberwachung führt aus technischen und strukturellen Gründen zu einer inhärent gegen das Prinzip der Datensparsamkeit verstoßenden, gesteigerten Erhebung und Verarbeitung von personenbeziehbaren Daten. Durch Verknüpfbarkeit der Kameras und Wiedererkennung von Personen ist nicht mehr nur Raumüberwachung sondern auch Personenüberwachung ermöglicht. Durch die Verknüpfbarkeit einzelner Systemen ist außerdem die Einhaltung der Zweckgebundenheit der Datenerhebung gefährdet. Durch eine kaum praktikable Einsichtnahme, Überprüfbarkeit der Datenspeicherung und -verarbeitung, sowie der Korrektur ist das Recht auf informationelle Selbstbestimmung stark eingeschränkt. Der gehobene Datenbedarf ergibt sich aus mehreren Tatsachen. Da Verhalten nicht eindeutig interpretiert werden kann, werden Systeme tendenziell misstrauisch konfiguriert, sodass Personen grundsätzlich intensiver beobachtet werden. Um die daraus resultieren-

de Fehlerrate zu reduzieren, müssen möglichst viele Daten erhoben und ausgewertet werden und auch Annahmen und Daten ausgewertet werden, die über das zu beobachtende Verhalten hinaus gehen. Auch weil im Vorhinein nicht klar ist, welche Daten für eine spätere Auswertung von Nutzen sein können, werden so viele Daten wie möglich erhoben. In wissenschaftlichen Publikationen vorgeschlagene Techniken zum Schutz der Identität im Bildmaterial sind nicht ausreichend zuverlässig und lassen bestimmte Identifizierungsmerkmale ungeschützt. Außerdem sind die Verfahren vorgesehen, um vor den Blicken der OperateurInnen zu schützen. Zur Beweissicherung bleiben die Daten weiterhin personenbeziehbar gespeichert.

Entgegen der allgemeinen Argumentation kann der Einsatz von OperateurInnen die umfangreiche Automatisierung nicht legitimieren und eine Verhinderung automatisierter Entscheidungen zum Nachteil Betroffener nicht sicherstellen. Für eine mündige Entscheidung können konzept- und technikbedingt keine adäquaten Informationen bereitgestellt werden. Durch die Übernahme des Systems des Filterns, Interpretierens und Schlussfolgerns stehen den OperateurInnen nur Informationen zur Verfügung, die stark durch die Sicht des Systems geprägt sind. Die Herleitung der Ergebnisse kann den OperateurInnen auf Grund der Komplexität des Systems nicht verständlich gemacht werden. Durch die Darstellungsweise der Informationen als integrierte dreidimensionale virtuelle Realität wird ein Eindruck von Vollständigkeit und Objektivität der Informationen vermittelt. Aus diesen Gründen ist der Entwurf einer eigenen Interpretation der Geschehnisse kaum möglich. Außerdem muss aus psychologischen Gründen mit einem übersteigerten Vertrauen gegenüber der Assistenz des Systems gerechnet werden. Zusätzlich zur ohnehin schon unzulänglichen Überprüfbarkeit führt dieses zu einer mangelnden Überprüfung der Darstellungen und Entscheidungen des System durch die

OperateurInnen. Diese kann nur teilweise und mit einem hohen Trainingsaufwand für OperateurInnen, der im Widerspruch zum Ziel der Kostensenkung steht, reduziert werden. Den OperateurInnen kann daher die Verantwortung für mögliche Fehler des Systems und Konsequenzen für Betroffene nicht übertragen werden.

Mit der grundsätzlichen Funktionsweise der Erkennung und Bewertung von Verhalten, geht außerdem die Gefahr einer diskriminierenden Wirkung einher. Einfallstore für diese liegen vor allem in der Gewinnung der Modelle, mit denen das erfasste Verhalten abgeglichen und ausgewertet wird. Sowohl bei manueller als auch bei automatisierter Gewinnung über Maschinenlernverfahren können intendiert oder ungewollt diskriminierende Aspekte in die Modelle gelangen oder bereits in den Algorithmen und Lernverfahren liegen. Diese Diskriminierung, wenn sie denn überhaupt identifiziert wird, kann auf Grund der komplexen Form in der die Modelle vorliegen, nur mit viel Aufwand oder gar nicht behoben werden. Der umfangreiche Einsatz der Systeme, Algorithmen oder Modelle kann zu einer Institutionalisierung der Diskriminierung führen, die über Rückkoppelungseffekte zu noch intensiveren Verfestigungen der Vorurteile führt.

Durch die massive Datenerhebung und die mangelnde Überprüfbarkeit entsteht eine, gegenüber *manueller Videoüberwachung* massiv gesteigerten Informationsasymmetrie. Während die Möglichkeiten der Erhebung, Verarbeitung, Nutzung und Speicherung enorm gesteigert sind, können Überwachte nicht nachvollziehen oder prüfen, ob und inwiefern diese Möglichkeiten ausgenutzt werden. Mit dem Wissen der Möglichkeiten über zeitlich rückwärts gerichtete Auswertung mit nachträglich festgelegten Kriterien hat automatisierte Videoüberwachung eine wesentlich intensivere selbstdisziplinierende Wirkung als *manuelle Videoüberwachung*. Ansätze, die Informationsasymmetrie durch Transparenz der Überwa-

chungsmaßnahme aufzubrechen, etwa durch Veröffentlichung der Quelltexte und der genauen Abläufe, sind aus technischen Gründen kaum umsetzbar und stehen im Konflikt mit dem wirtschaftlichen Interesse der Hersteller und dem Sicherheitskonzept der Betreiber. Selbst mit diesen Informationen kann die tatsächliche Funktionsweise der Systeme nicht zweifelsfrei überprüft werden.

Der Einsatz automatisierter Videoüberwachung hat durch einzelne Maßnahmen nicht nur unmittelbare Auswirkungen auf Individuen, sondern schafft durch die Tatsache, dass automatisiert überwacht wird, eine Stimmung des Misstrauens. Auf diese Weise haben selbst *manuelle Videoüberwachung*sanlagen, oder weniger umfangreich automatisierte Videoüberwachung den gleichen disziplinierenden Effekt auf Menschen, da die Verarbeitung nicht ausgeschlossen werden kann. Selbst wenn jede einzelne Maßnahme in ihrer Ausprägung vollkommen mit dem Recht vereinbar wäre, erzeugen die Kameras den Eindruck einer flächendeckenden Überwachung und erzeugen eine Stimmung des Misstrauens und des Verdachtes. Automatisierte Videoüberwachung kann daher eine positive gesellschaftliche Entwicklung gefährden.

Von bestimmten Interessengruppen kann dieser umfangreiche disziplinierende Effekt sogar erwünscht sein. Die Technik, die eine umfangreiche, unspezifische, tiefgreifende, flächendeckende und zeitlich unabhängige Überwachung ohne großen personellen Aufwand ermöglicht, birgt die Gefahr als Werkzeug der Unterdrückung missbraucht zu werden.

6.2 Fazit

Wenn die Sphäre des Herstellens in den Raum wesentlichen Handelns eingedrungen ist, dann muss Moralität in die Sphäre des Herstellens eindringen.[324]

HANS JONAS

Technik hat eine in der Geschichte der Menschheit beispiellose sowohl zeitliche als auch räumliche kausale Reichweite.[325] Besonders Informationstechnik dringt in immer mehr Bereiche des gesellschaftlichen und des individuellen alltäglichen Lebens vor. Probleme solcher Techniken wirken sich dementsprechend immer stärker auf Individuen und Gesellschaft aus, ohne, dass diese Folgen für EntwicklerInnen und NutzerInnen der Techniken zwingend direkt mit ihr in Verbindung gebracht oder vorausgesehen werden. Die Geschichte zeigte, dass umgesetzt wird, was technisch machbar ist. Am Beispiel der Videoüberwachung wurde deutlich, dass weder die Effektivität nachgewiesen noch Gefahren untersucht bzw. diese einfach ignoriert wurden und der Ausbau massiv vorangetrieben wurde.

Nach Hans Jonas muss jedoch eben diese „vorausgedachte Gefahr" als Kompass für eine ethische Bewertung dienen. Erst durch das Vorausdenken der Probleme können Reglementierungen zum Schutz der Menschen- und Grundrechte ausgestaltet werden, die Verhältnismäßigkeit von Maßnahmen ausreichend akkurat eingeschätzt werden, oder Probleme schon während der Entwicklung berücksichtigt werden.

Mit der Diplomarbeit und in diesem Buch wurde eben dieser Versuch des Vorausdenkens der Gefahr für die Automatisierung von Videoüberwachung unternommen – ohne dabei jedoch Anspruch auf Vollständigkeit zu erheben. Ergebnis sind nicht nur die identifizierten Probleme selbst.

---

324 **Jonas**, *Das Prinzip Verantwortung : Versuch einer Ethik für die technologische Zivilisation*, S. 32.

325 Vgl. ebd., S. 22.

Aus der Art der Probleme und dem Vorgehen zur Identifizierung können Implikationen für Herangehensweise von Technikfolgeabschätzung im Bereich der Informationstechnik und die Verantwortung der Informatik als Wissenschaft abgeleitet werden.

6.2.1 Verantwortung der Informatik

Es konnte gezeigt werden, dass zum Teil subtile technische Funktionsweisen automatisierter Videoüberwachung zu negativen Auswirkungen mit großer gesellschaftlicher Tragweite führen können. Da nur die Informatik mit ihrem technischen Wissen in der Lage ist, derartige Probleme zu identifizieren, hat die Informatik als Wissenschaft eine große Verantwortung inne. Um dieser umfassend nachzukommen, muss eine Sensibilität für gesellschaftliche und ethische Aspekte elementarer Bestandteil von Forschung und Lehre der Informatik werden. Dieses Buch zeigt jedoch auch, dass allein technisches Wissen für eine Abschätzung nicht genügt. Wissen aus anderen Wissenschaftsbereichen – hier vor allem Soziologie und Psychologie – müssen mit einbezogen werden und daraufhin die Technik untersucht werden.

Das Buch macht außerdem darauf aufmerksam, dass der überwachungskritische Diskurs sich fast ausschließlich auf die Datenschutzproblematik bezieht und andere gesellschaftliche Auswirkungen und Probleme weitestgehend außer Acht gelassen werden. Dies birgt die Gefahr, dass Videoüberwachung als verhältnismäßiger oder sogar unproblematisch eingeschätzt wird, wenn Techniken eingesetzt werden, die den Datenschutzansprüchen vermeintlich gerecht werden. Tatsächlich konnte jedoch gezeigt werden, dass bezüglich automatisierter Videoüberwachung weder Datenschutztechniken noch der Einsatz von Menschen, Mängel und Probleme ausgleichen können. Es wurde mehrfach der generell zu beobachtende Technikglaube angesprochen. Die Verantwortung der Informatik

liegt daher auch in der Aufklärung und der Richtigstellung von falschen Vorstellungen und der aktiven Bereicherung des Diskurses.

6.2.2 Verhältnismäßigkeit automatisierter Videoüberwachung

Das entscheidendste Resultat ist jedoch die Erkenntnis, dass die Verhältnismäßigkeit von Effektivität und resultierenden Einschränkungen nicht abgeschätzt werden kann. Probleme konnten zwar identifiziert werden, *wie* tiefgreifend und *wie* umfassend sich diese tatsächlich auf Individuen und Gesellschaft auswirken, kann jedoch nicht vorausgedacht werden. Obwohl mangels Quellen nicht explizit auf Effektivität eingegangen werden konnte, wurden doch grundsätzliche, inhaltliche und auch technische Probleme genannt, die begründete Zweifel an einer Effektivität aufkommen lassen. Zusammenfassend spricht also vieles gegen eine hohe Effektivität und vieles für gravierende Einschränkungen.
Es wäre verwerflich, allein auf Grund des *Glaubens an das Konzept Videoüberwachung*, ohne Nachweis der Effektivität unter Ignoranz der Gefahren, derartige Maßnahmen zu fordern. Es könnte verheerend sein, derartige Überwachungsmaßnahmen zu installieren.
Von Automatisierung der Videoüberwachung und Vernetzung bestehender manueller Überwachungsmaßnahmen sollte grundsätzlich abgesehen werden, da sogar unabhängig von einzelnen Maßnahmen durch die generelle Tatsache, dass automatisierte Videoüberwachung überhaupt stattfindet, Demokratie und gesellschaftliche Entwicklung gefährdet wird.
Der Forschung und der Kommerzialisierung derartiger Technik für privaten und öffentlichen Gebrauch sollten enge Grenzen gesetzt werden. Projekte wie INDECT, deren Forschungsresultate im Einsatz eindeutig nicht mit europäischem Recht zu vereinbaren wären, müssen unterbunden werden. Der Export derartiger Systeme oder Teilsysteme muss verhindert werden.

Literatur

Adorno, Theodor: *Gesammelte Schriften : Kulturkritik und Gesellschaft II*. Frankfurt am Main: Suhrkamp, 1977.

Albrecht, Peter-Alexis: *Der Weg in die Sicherheitsgesellschaft : Auf der Suche nach staatskritischen Absolutheitsregeln*. 1. Auflage, Berlin: BWV Verlag, 2010.

Amicelle, Anthony: *Exclusion and discrimination*. In: *Deliverable D1.1: Surveillance, fighting crime and violence.* IRISS : Increasing Resilience in Surveillance Societies, 2012, S. 220–226.

Apelt, Maja; **Möllers**, Norma: *Wie „intelligente" Videoüberwachung erforschen? : Ein Resümee aus zehn Jahren Forschung zu Videoüberwachung*. In: *Zeitschrift für Außen- und Sicherheitspolitik*. 4 VS-Verlag, 2011, S. 585–593.

Bahner, Jennifer: *Übersteigertes Vertrauen in Automation: der Einfluss von Fehlererfahrungen auf complacency und Automation-Bias*. Diss. Berlin Institute of Technology, 2008.

Baiget, Pau: *Observing Human Behavior in Image Sequences: the Video-Hermeneutics Challenge*. In: *Computer Vision : Advances in Research and Development (CVCRD 2009).* Bellaterra, Spanien, 2008.

Ball, Kirstie: *surveillancy and conformity*. In: *Deliverable D1.1: Surveillance, fighting crime and violence.* IRISS : Increasing Resilience in Surveillance Societies, 2012, S. 226–233.

BMBF (Hrsg.): *Forschung für die zivile Sicherheit 2012 – 2017 : Rahmenprogramm der Bundesregierung*. Bonn, 2012.

Bodor, Robert: *Optimal Camera Placement for Automated Surveillance Tasks*. In: *Journal of Intelligent and Robotic Systems*. 50.3, Springer Netherlands, 2007, S. 257–295.

Bygrave, Lee: *Minding the machine: Article 15 of the EC Data protection Directive and Automated Profiling*. In: *EC Data protection Directive and Automated Profiling*. Computer Law & Security Report, 2001, S. 17–24.

Cannataci, Joseph: *Squaring the Circle of Smart Surveillance and Privacy*. In: *Digital Society, 2010. ICDS '10. Fourth International Conference on.* 2010, S. 323–328.

Coudert, Fanny; **Dumortier**, Jos: *Intelligent Video Surveillance Networks: Data Protection Challenges*. In: *2012 Seventh International Conference on Availability, Reliability and Security*. 0 Los Alamitos, CA, USA: IEEE Computer Society, 2008, S. 975–981.

Cristani, Marco: *Look at Who's Talking: Voice Activity Detection by Automated Gesture Analysis*. In: *AmI Workshops.* 2011, S. 72–80.

Dao, Minh-Son: *Abandoned Object's Owner Detection: A Case Study of Hybrid Mobile-Fixed Video Surveillance System*. In: *Advanced Video and Signal Based Surveillance, IEEE Conference on*. 0 Los Alamitos, CA, USA: IEEE Computer Society, 2012, S. 404–409.

DeCamp, Philip: *An immersive system for browsing and visualizing surveillance video*. In: *Proceedings of the international conference on Multimedia.* MM '10 New York, NY, USA: ACM, 2010, S. 371–380.

Dee, Hannah; **Velastin**, Sergio: *How close are we to solving the problem of automated visual surveillance?* In: *Machine Vision and Applications*. 19 Springer-Verlag, 2008, S. 329–343.

Deutscher Bundestag (Hrsg.): *Drucksache 17/2750*. Berlin, 2010.

Deutscher Bundestag (Hrsg.): *Drucksache 17/3940*. Berlin, 2010.

Dix, Alexander: *Datenschutz und Informationsfreiheit : Bericht 2010*. In: *Datenschutz und Informationsfreiheit in Berlin*. 2011.

Dufaux, Frederic; **Ebrahimi**, Touradj: *A framework for the validation of privacy protection solutions in video surveillance*. In: *2012 IEEE International Conference on Multimedia and Expo*. o Los Alamitos, CA, USA: IEEE Computer Society, 2010, S. 66–71.

Dziech, Andrzej: *INDECT*. In: *Security R&D Innovation for the Citizens.* Stockholm: INDECT, 2009.

d'Angelo, David: *CamInSens : An Intelligent in-situ Security System for Public Spaces*. In: *International Conference on Security and Management (SAM)*. Fraunhofer IAIS ; Leibniz University Hannover, 2012.

Erhardt, Angelika: *Einführung in die Digitale Bildverarbeitung : Grundlagen, Systeme und Anwendungen*. Berlin, Heidelberg: Springer, 2007.

Eugster, Hannes; **Nebiker**, Stephan: *UAV-Based augmented monitoring : real-time georeferencing and integration of video imagery with virtual globes*. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B1. Beijing, 2008, S. 1229–1236.

Europäische Kommision (Hrsg.): *Vorschlag für Verordnung des europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr (Datenschutz-Grundverordnung)*. Brüssel, 2012.

Fidis WP6 (Hrsg.): *D6.7c: Forensic Profiling*. Techn. Ber. FIDIS – Future of Identity in the Information Society, 2008.

Galliker, Mark; **Wagner**, Franc: *Ein Kategoriensystem zur Wahrnehmung und Kodierung sprachlicher Diskriminierung*. In: *Journal für Psychologie*. 3.3, 1995, S. 33–43.

Gehring, Dirk; **Kühne**, Hildergard; **Schultz**, Tanja: *Erkennung von menschlichen Bewegungen mit Hidden Markov Modellen*. In: *dvs Band*. Nomos, 2011.

Green, David; **Swets**, John: *Signal Detection Theory and Psychophysics*. Peninsula Publishing, 1988.

Grunwald, Armin: *Technikfolgenabschätzung: eine Einführung*. Gesellschaft – Technik – Umwelt Edition Sigma, 2010.

Hempel, Leon; **Alisch**, Christian: *Evaluation der 24-Stunden-Aufzeichnung in Ubahnstationen der Berliner Verkehrsbetriebe (BVG) : Zwischenbericht*. D:4 BaSE – Büro für angewandte Statistik und Evaluation, 2006.

Hempel, Leon; **Töpfer**, Eric: *Videoüberwachung in Europa : Abschlussbericht*. In: *Urbaneye Arbeitspapier Nr. 15*. 2004.

Hu, Weiming: *A survey on visual surveillance of object motion and behaviors*. In: *IEEE Transactions on Systems, Man and Cybernetics*. 34 2004, S. 334–352.

Hähner, Jörg; **Grenz**, Carsten; **Jänen**, Uwe: *Verteilte vernetzte Kamerasysteme zur in-situ Erkennung Personen-induzierter Gefahrensituationen*. Techn. Ber. CamInSens, 2012.

Introna, Lucas; **Nissenbaum**, Helen: *The Internet as a democratic medium: why the politics of search engines matters*. In: *Information Society 16(3)*. 2000, S. 169–185.

Introna, Lucas; **Wood**, David: *Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems*. In: *Surveillance & Society 2(2/3)*. 2004, S. 177–198.

Jonas, Hans: *Das Prinzip Verantwortung : Versuch einer Ethik für die technologische Zivilisation*. 4. Auflage, Berlin: Suhrkamp Verlag GmbH, 2003.

Khalid, Shehzad: *Motion-based behaviour learning, profiling and classification in the presence of anomalies*. In: *Pattern Recogn*. 43.1, New York, NY, USA: Elsevier Science Inc., 2010, S. 173–186.

Klapaftis, Ioannis; **Manandhar**, Suresh; **Pandey**, Shailesh: *XML Data Corpus: Report on methodology for collection, Deliverable name cleaning and unified representation of large textual data from various sources: news reports, weblogs, chat*. 2009.

Klauser, Francisco: *Die Videoüberwachung öffentlicher Räume : Zur Ambivalenz eines Instruments sozialer Kontrolle*. 1. Auflage, Frankfurt am Main: Campus Verlag, 2006.

Ko, Teddy: *A Survey on Behavior Analysis in Video Surveillance Applications*. In: *Video Surveillance*. InTech, 2011, S. 279–294.

Kreissl, Reinhard: *The effectiveness of surveillance in preventing and detecting crime and terrorism*. In: *Deliverable D1.1: Surveillance, fighting crime and violence.* IRISS : Increasing Resilience in Surveillance Societies, 2012, S. 159–213.

Macnish, Kevin: *Surveillance Ethics*. In: *The Internet Encyclopedia of Philosophy*. The Internet Encyclopedia of Philosophy, 2011.

Macnish, Kevin: *Unblinking eyes : the ethics of automating surveillance*. In: *Ethics and Information Technology*. 14.2, Springer Netherlands, 2012, S. 151–167.

Manzey, Dietrich; **Bahner**, Jennifer: *Vertrauen in Automation als Aspekt der Verllässlichkeit von Mensch-Maschine-Systemen*. In: *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis: Festschrift für Klaus-Peter Timpe*. Symposion Publishing GmbH, 2005.

Moncrieff, Simon; **Venkatesh**, Svetha; **West**, Geoff: *Dynamic Privacy in Public Surveillance*. In: *Computer*. 42.9, Los Alamitos, CA, USA: IEEE Computer Society, 2009, S. 22–28.

Mosier, Kathleen; **Skitka**, Linda: *Human Decision Makers and Automated Decision Aids: Made for Each Other?* In: *Automation and Human Performance: Theory and Applications*. Lawrence Erlbaum Assoc Inc, 1996, S. 201–220.

Mosier, Kathleen; **Skitka**, Linda; **Burdick**, Mark: *Automation bias: Decision making and performance in high-tech cockpits*. In: *International Journal of Aviation Psychology*. 8.1, USA: San Jose State University ; NASA Ames Research Center Moffett Field, 1997, S. 47–63.

Musik, Christoph: *The thinking eye is only half the story: high-level semantic video surveillance*. In: *Information Polity*. 16.4, Amsterdam, Niederlande: IOS Press, 2011, S. 339–353.

Möllers, Norma; **Hälterlein**, Jens: *Privacy issues in public discourse: the case of "smart" CCTV in Germany*. In: *Innovation: The European Journal of Social Science Research*. 2012, S. 1–14.

Newton, Elaine; **Sweeney**, Latanya; **Malin**, Bradley: *Preserving Privacy by De-Identifying Face Images*. In: *IEEE Transactions on Knowledge and Data Engineering*. 17.2, Piscataway, NJ, USA: IEEE Educational Activities Department, 2005, S. 232–243.

Norris, Clive; **Armstrong**, Gary: *CCTV and the social Structuring of Surveillance*. In: *Crime Prevention Studies Series*. 10 Criminal Justice Press, 1999, S. 157–178.

Raab, Charles: *Impact of Surveillance on civil liberties and fundamental rights*. In: *Deliverable D1.1: Surveillance, fighting crime and violence.* IRISS : Increasing Resilience in Surveillance Societies, 2012, S. 254–302.

Robertson, Neil; **Reid**, Ian; **Brady**, Michael: *Automatic human behaviour recognition and explanation for CCTV video surveillance*. In: *Security Journal*. 2007.

Roßnagel, Alexander; **Desoi**, Monika; **Hornung**, Gerrit: *Gestufte Kontrolle bei Videoüberwachungsanlagen : Ein Drei-Stufen-Modell als Vorschlag zur grundrechtsschonenden Gestaltung*. In: *Datenschutz und Datensicherheit*. 35.10, 2011, S. 694–701.

Rothmann, Robert: *Zur Evaluation der Sicherheitstechnischen Eignung von Videoüberwachung. Regionale Defizite, internationale Standards, methodische Herausforderungen*. In: *juridikum, zeitschrift für kritik | recht | gesellschaft*. 4/2012 2012, S. 481–493.

Ruegg, Jean; **November**, Valérie; **Klauser**, Francisco: *CCTV, risk management and regulation mechanisms in publicly-used places : a discussion based on Swiss examples*. In: *Surveillance & society*. 2.2/3, Surveillance & Society, 2004, S. 415–429.

Röbke, Oliver: *Automatisierte Detektion interventionsbedürftiger Situationen durch Klassifizierung visueller Muster (ADIS)*. Techn. Ber. Bundesministerium für Bildung und Forschung, 2012.

Satta, Riccardo; **Fumera**, Giorgio; **Roli**, Fabio: *Fast person re-identification based on dissimilarity representations*. In: *Pattern Recognition Letters*. 33.14, 2012, S. 1838 –1848.

Scheuer, Josef: *Supporting Video Surveillance by Computer Graphics*. Magisterarb. TU-Delft, 2007.

Scholz, Philip: § *6b 2.2*. In: *BDSG*. Nomos, 2011.

Schwabach, Helmut: *Distributed Embedded Smart Cameras for Surveillance Applications*. In: *IEEE Computer*. 39.2, 2006, S. 68–75.

Senior, Andrew: *Protecting Privacy in Video Surveillance*. Berlin, Heidelberg: Springer London, 2009.

Senior, Andrew: *Enabling Video Privacy through Computer Vision*. In: *IEEE Security & Privacy*. 3.3, Los Alamitos, CA, USA: IEEE Computer Society, 2005, S. 50–57.

Sester, Monika; **Kuntzsch**, Colin: *Szenenanalyse – Mustererkennung in Personen-Tracks*. Techn. Ber. Institut für Kartographie und Geoinformatik, 2013.

Thiel, Geoff: *Automatic CCTV surveillance-towards the VIRTUAL GUARD*. In: *Aerospace and Electronic Systems Magazine, IEEE*. 15.7, 2000, S. 3–9.

Tichy, Gunther; **Peissl**, Walter: *Beeinträchtigung der Privatsphäre in der Informationsgesellschaft*. Techn. Ber. Institute of Technology Assessment (ITA), 2001.

Töpfer, Eric: *Videoüberwachung als Kriminalprävention? Plädoyer für einen Blickwechsel*. In: *Kriminologisches Journal*. Jg. 41, Nr. 4 Hamburg, 2009, S. 272–282.

Valera, Maria; **Velastin**, Sergio: *Intelligent distributed surveillance systems: A review*. In: *IEEE Proceedings : Vision, Image and Signal Processing*. 2005.

Wandke, Hartmut; **Wetzenstein**, Elke; **Polkehn**, Knut: *Handlungsbezogene Elementarbausteine für Fahrerassistenzsysteme*. In: *Der Fahrer im 21. Jahrhundert. VDI-Berichte 1919*. Düsseldorf: VDI-Verlag, 2005, S. 41–62.

Weichert, Thilo: *Private Videoüberwachung und Datenschutzrecht*. In: *Detektiv-Kurier*. 04/2001 2001.

Welsh, Brandon; **Farrington**, David: *Crime Prevention Effects of Closed Circuit Television: A Systematic Review*. Home Office Research studies Home Office Research, Development ; Statistics Directorate, 2002.

Westin, Alan: *Privacy and Freedom*. London: Bodley Head, 1970.

Wiliem, Arnold: *A Context-Based Approach for Detecting Suspicious Behaviours*. In: *Proceedings of the 2009 Digital Image Computing: Techniques and Applications*. DICTA '09 Washington, DC, USA: IEEE Computer Society, 2009, S. 146–153.

Wiliem, Arnold: *A Context Space Model for Detecting Anomalous Behaviour in Video Surveillance*. In: *Proceedings of the 2012 Ninth International Conference on Information Technology : New Generations*. ITNG '12 Washington, DC, USA: IEEE Computer Society, 2012, S. 18–24.

Winkler, Thomas: *Vertrauenswürdige Videoüberwachung : Sichere intelligente Kameras mit Trusted Computing*. In: *Datenschutz und Datensicherheit*. 35.11, 2011, S. 797–801.

Winner, Langdon: *Do Artifacts Have Politics?* In: *Chicago, University of Chicago Press*. 1986, S. 19–39.

Xudong, Zhu; **Hui**, Li; **Zhijing**, Liu: *Behavior Clustering for Anomaly Detection*. In: *China Communications*. 7.6, China Communications, 2010, S. 17.

Yingjie, Li; **Yin**, Yixin: *Towards Suspicious Behavior Discovery in VideoSurveillance System*. In: *Second International Workshop on Knowledge Discovery and Data Mining*. IEEE Computer Society, 2009, S. 539–541.

Čas, Johann: *The relevance of social and economic costs of surveillancy – Conclusion*. In: *Deliverable D1.1: Surveillance, fighting crime and violence*. IRISS : Increasing Resilience in Surveillance Societies, 2012, S. 252–253.

Onlinequellen

3Sat (Hrsg.): *Kulturzeit : INDECT*. YouTube 2011 URL: `https://www.youtube.com/watch?v=F_izSRnT98Q` [Stand 25.08.2012].

Beaupoil, André: *Das Netzwerk der NSA : Horchposten – auch in Deutschland*. 2013 URL: `http://www.tagesschau.de/nsa-video100.html` [Stand 25.08.2012].

Becker, Andreas: *Automatische Absichtserkennung*. 2013 URL: `http://www.heise.de/tp/artikel/36/36630/1.html` [Stand 25.08.2012].

Behörden Spiegel (Hrsg.): *Mehr als nur Kamera-Überwachung*. 2013 URL: `http://www.behoerden-spiegel.de/icc/Internet/sub/94c/94c34370-af6f-c311-994a-0927b988f2ee,,,aaaaaaaa-aaaa-aaaa-bbbb-000000000011` [Stand 25.08.2012].

BeWare (Hrsg.): *Business case and example scenario*. YouTube 2010 URL: `http://www.youtube.com/watch?v=-LM_4d9ev-k` [Stand 25.08.2012].

Carroll, Rory: *Welcome to Utah, the NSA's desert home for eavesdropping on America*. 2013 URL: http://www.guardian.co.uk/world/2013/jun/14/nsa-utah-data-facility [Stand 25.08.2012].

Clifford, Stephanie; **Hardy**, Quentin: *Attention, Shoppers: Store Is Tracking Your Cell*. 2013 URL: http://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html?_r=0 [Stand 25.08.2012].

Coherent Synchro, *Coherent Synchro 3D Visualization Platform*. YouTube 2013 URL: https://www.youtube.com/watch?v=Xf1_YP0fQSo [Stand 25.08.2012].

Cuxhavener Nachrichten (Hrsg.): *Die „kleinen Brüder" schauen nie weg: Kamera-Überwachung in Cuxhaven*. 2004 URL: http://www.cn-online.de/lokales/news/die-kleinen-brueder-schauen-nie-weg-kamera-ueberwachung-in-cuxhaven.html [Stand 25.08.2012].

Erling, Johnny: *Chinas Polizei will den totalen Überblick*. 2009 URL: http://www.welt.de/politik/ausland/article4308754/ [Stand 25.08.2012].

Frontal 21, *Sendung vom 26.01.2010*. ZDF TV 2010 URL: http://www.youtube.com/watch?v=5x7J9kq86Qc [Stand 25.08.2012].

Futurezone.at (Hrsg.): *Deutsche Bahn will Drohnen gegen Sprayer*. 2013 URL: http://futurezone.at/digitallife/16121-deutsche-bahn-will-drohnen-gegen-sprayer.php [Stand 25.08.2012].

INDECT, *INDECT Präsentationsfilm*. YouTube 2011 URL: http://www.youtube.com/watch?v=9gVBFJg1AbA [Stand 25.08.2012].

INDECT (Hrsg.): *INDECT FAQ: Frequently Asked Questions*. 2012 URL: http://www.indect-project.eu/faq [Stand 25.08.2012].

Kannenberg, Axel: *Bayerische Datenschützer: Schon 17.000 kommunale Überwachungskameras*. 2013 URL: http://www.heise.de/-1913276.html [Stand 25.08.2012].

Kleinz, Torsten: *Generalverdacht der Algorithmen*. 2013 URL: http://notes.computernotizen.de/2013/06/19/generalverdacht-der-algorithmen/ [Stand 25.08.2012].

Krempl, Stefan: *Berlin will Videoüberwachung mit biometrischer Gesichtserkennung testen*. 2008 URL: http://www.heise.de/-204147.html.

Meyer, Angela: *Zweifelhafter Notanker : Videoüberwachung in Schulen*. 2004 URL: http://heise.de/-289218 [Stand 25.08.2012].

Monroy, Matthias: *Allround-System für europäische Homeland Security*. 2010 URL: http://www.heise.de/tp/r4/artikel/31/31802/1.html [Stand 25.08.2012].

Monroy, Matthias: *Mehr Polizeidrohnen im Anflug*. 2011 URL: http://www.heise.de/tp/artikel/34/34202/1.html [Stand 25.08.2012].

NEXCOMInternational (Hrsg.): *Intelligent Surveillance Solutions*. YouTube 2011 URL: http://www.youtube.com/watch?v=ey8r0A_5B7Q [Stand 25.08.2012].

Nogala, Detlef: *Der Frosch im heißen Wasser : Die Trivialisierung von Überwachung in der informatisierten Gesellschaft des 21. Jahrhunderts*. 2000 URL: http://www.heise.de/tp/artikel/8/8988/1.html [Stand 25.08.2012].

Poitras, Laura: *PRISM Whistleblower: Hong Kong*. Praxis Films – YouTube 2013 URL: http://www.youtube.com/watch?v=5yB3n9fu-rM [Stand 25.08.2012].

Reeve, Tom: *BSIA attempts to clarify question of how many CCTV cameras there are in the UK*. 2013 URL: http://www.securitynewsdesk.com/2013/07/11/bsia-attempts-to-clarify-question-of-how-many-cctv-cameras-in-the-uk/ [Stand 25.08.2012].

Rutz, Charlie: *Interview mit INDECT-Projekt-Koordinator*. 2011 URL: http://freidenker.cc/wp-content/uploads/Interview_mit_INDECT-Projekt-Koordinator.pdf [Stand 25.08.2012].

Spool, Jared: *The $300 Million Button*. 2009 URL: https://www.uie.com/articles/three_hund_million_button/ [Stand 25.08.2012].

Suarez, Daniel: *The kill decision shouldn't belong to a robot*. TedTalk Video 2010 URL: http://on.ted.com/DSuarez [Stand 25.08.2012].

Wahlbrink, Joachim: *Zahlreiche Rechtsverstöße bei der Videoüberwachung : Wahlbrink: Behörden und Kommunen ignorieren Datenschutzgesetz*. 2010 URL: http://www.lfd.niedersachsen.de/portal/live.php?navigation_id=13110 [Stand 25.08.2012].

Wilkens, Andreas: *PRISM-Überwachung : BND und NSA in einem Boot*. 2013 URL: http://www.heise.de/-1917850.html [Stand 25.08.2012].

Zeit Online (Hrsg.): *Verräterisches Handy*. 2010 URL: http://www.zeit.de/datenschutz/malte-spitz-vorratsdaten [Stand 25.08.2012].

Abbildungsverzeichnis