

Intelligente Videoüberwachung

MIT INFORMATIONSTECHNOLOGIE UND SOZIOLOGISCHEN ERKENNTNISSEN
DEN ALLTAG SICHERER MACHEN

Videoüberwachung von öffentlichen Plätzen ist eine mittlerweile alltägliche Erscheinung. Im BMBF-geförderten Projekt »Automatische Situations-einschätzung für Ereignis-gesteuerte Videoüberwachung« (ASEV) forschen Wissenschaftler aus fünf Instituten der Leibniz Universität Hannover daran, eine Technologie zu entwickeln, die sowohl die Effizienz der Operateure durch eine automatische Vorauswertung der Bilder erhöht als auch den Missbrauch der Videodaten verhindert. Verwendet werden dafür intelligente Verfahren der Bilderkennung und Erkenntnisse aus der Soziologie.

Das Problem mit der Videoüberwachung

Die Sicherheit auf öffentlichen Plätzen, in Industrieanlagen, auf Flughäfen oder in U-Bahn-Stationen hängt stark von Überwachungssystemen ab, bei denen geschultes Personal (sogenannte Operateure) Videobilder auswerten und einschätzen.

Diese Form der Videoüberwachung hat allerdings schwerwiegende Nachteile: Zum einen ist die Überwachung von Videobildern für das Sicherheitspersonal sehr ermüdend und damit stark fehleranfällig. So hat sich in Studien herausgestellt, dass eine aufmerksame visuelle Auswertung nur etwa zwanzig Minuten lang ohne Pause möglich ist. Das andere Problem betrifft den Datenschutz und den Schutz von Persönlichkeitsrechten bei der Videoüberwachung: In allen oben genannten Szenarien muss gewährleistet werden, dass die Kamerabilder nur wie vorgesehen verwendet und nicht zur Mitarbeiterüberwachung oder für Stalking missbraucht werden.

ASEV – automatisierte Videoüberwachung

Im Projekt ASEV wird ein intelligentes Bilderkennungssystem für Flughäfen entwickelt, das eine computergestützte

Am Projekt beteiligte Institute

- Institut für Informationsverarbeitung: *Echtzeitfähige Merkmalsverarbeitung*
- Forschungszentrum L3S: *Automatisches Einschätzen und Erklären von Situationen*
- Institut für Mikroelektronische Systeme – Fachgebiet Architekturen und Systeme: *Hardware-unterstützte Bildauswertung*
- Regionales Rechenzentrum für Niedersachsen: *Gesicherte Datenübertragung und -speicherung*
- Institut für Soziologie: *Untersuchung gesellschaftlicher und rechtlicher Aspekte der automatisierten Situations-einschätzung*

Vorauswertung durchführt. Dieses System soll die Operateure bei ihrer Arbeit unterstützen, so dass Fehleinschätzungen vermieden werden. ASEV wird vom Bundesministerium für Bildung und Forschung (BMBF) gefördert und hat neben den wissenschaftlichen Partnern der Leibniz Universität auch Firmen aus dem Bereich Sicherheitssysteme

und Kameraüberwachung sowie als assoziierte Partner die Flughäfen Braunschweig, Hannover und Hamburg an Bord. Ziel von ASEV ist es, Grundlagen für die automatisierte Erkennung von gefährlichen oder potenziell gefährlichen Situationen auf Flughäfen zu erforschen und ein sozial verträgliches System aus Hard- und Software zu entwickeln, das die Flughafensicherheit erhöht.

Videoanalyse durch Merkmalsvergleich

Die Fähigkeit, komplexe Bilder zu analysieren und Objekte zu erkennen, muss dem Computer durch vorgegebene Verarbeitungsschritte ermöglicht werden, denn dies kann mit Künstlicher Intelligenz bislang nicht selbstständig erlernt werden. Ein aktuelles Verfahren ist, aus einem gegebenen Bild Merkmale zu extrahieren. Ein Merkmal besteht aus der Merkmalsposition und dem Deskriptor – üblicherweise ein hochdimensionaler Vektor mit einer Reihe von Werten, die aus dem lokalen Bildinhalt ermittelt werden. Die Idee: Sind sich zwei Deskriptoren ähnlich, so sehen Merkmale ähnlich aus. Stimmen viele Merkmale überein, so handelt es sich um ähnlich aussehende Objekte. Das Problem, Bilder zu analysieren, wird dadurch vereinfacht auf die Aufgabe,

(sehr viele) Vektoren zu vergleichen.

In ASEV werden zur Objektbeschreibung sogenannte SIFT-Merkmale (Scale Invariant Feature Transform) eingesetzt. Bei diesem Verfahren werden automatisch besonders gut wiederzufindende Bildbereiche detektiert. Anschließend wird der Deskriptor bewusst so gebildet, dass eine Änderung der Beleuchtung im Bild sowie Skalierung oder Rotation keinen Einfluss haben. Am

Zeitpunkten und Kameras zu fusionieren.

Das Ergebnis eines Merkmalsvergleichs ist eine Liste von Punktübereinstimmungen (Korrespondenzen) zwischen mehreren Bildern. Da Korrespondenzen den mathematischen Abbildungsgesetzen der Kamera gehorchen müssen, können ungültige Korrespondenzen gefiltert und somit »zufällige« Übereinstimmungen ausgeschlossen werden. Nach diesem Schritt liegen also gesicherte Informationen

Bildverarbeitungsaufgaben schon direkt nach der Aufnahme durchzuführen, besonders dann, wenn das System modular und skalierbar ausgelegt sein soll. Die Idee ist, dass die Kamera als sogenannte »Smart Camera« ausgeführt wird, das heißt, dass die Kamera statt der Rohbilddaten ein bereits ausgewertetes Bild mit Zusatzinformationen liefert. Solche Zusatzinformationen wären für das ASEV-Projekt zum Beispiel im Bild wiedererkannte Objekte. Ermöglicht wird das durch spezielle anpassbare Prozessoren (ASIPs = Application Specific Instruction Set Processors), die am Lfi durch das Institut für Mikroelektronische Systeme (IMS) für die genaueren Bedürfnisse der

Abbildung 1 Die detaillierte Analyse des Bildinhaltes ist Voraussetzung für eine Situationseinschätzung. Quelle: Institut für Informationsverarbeitung



onstechnologie (Lfi) wird vom Institut für Informationsverarbeitung (TNT) zusätzlich an einer Verbesserung der perspektivischen Invarianz geforscht.

Um gezielt Objekte zu erkennen, müssen zunächst Objektmerkmale gelernt werden, was bei dem SIFT-Ansatz lediglich das Abspeichern aller Merkmale bedeutet. Schwierig wird es allerdings dadurch, dass nicht alle Merkmale gleichzeitig zu sehen sind. Am TNT werden deswegen Algorithmen entwickelt, um Merkmale aus verschiedenen Objektsichten von verschiedenen

über Objekte in den Kamerabildern vor.

Gesucht: Genügend Rechenleistung

Das geschilderte Verfahren erfordert allerdings eine sehr hohe Rechenleistung, besonders, wenn viele hochauflösende Kameras eingesetzt werden. Prinzipiell ließe sich die Rechenleistung durch ein geeignet dimensioniertes zentrales Rechnersystem bereitstellen, was für viele Anwendungen jedoch inpraktikabel ist. Ein interessanter Ansatz ist es daher, möglichst viele der

angewendeten Algorithmen entworfen und getestet werden.

Sowohl General-Purpose-Prozessoren als auch ASIPs besitzen als Grundvokabular einen Satz an Instruktionen, wie zum Beispiel Addition, Subtraktion, Multiplikation oder logische Verknüpfungen. General-Purpose-Prozessoren können nur Instruktionen bearbeiten, die für möglichst viele Aufgaben geeignet sind. Komplexere Rechenoperationen, wie zum Beispiel die Division, für die keine Instruktion bereitsteht, müssen in mehrere verfüg-

Abbildung 2 Im Gegensatz zur traditionellen manuellen Analyse ermöglicht eine automatische Situationseinschätzung eine genaue Auswertung bei gleichzeitiger Wahrung der Privatsphäre. Quelle: iStockphoto.com

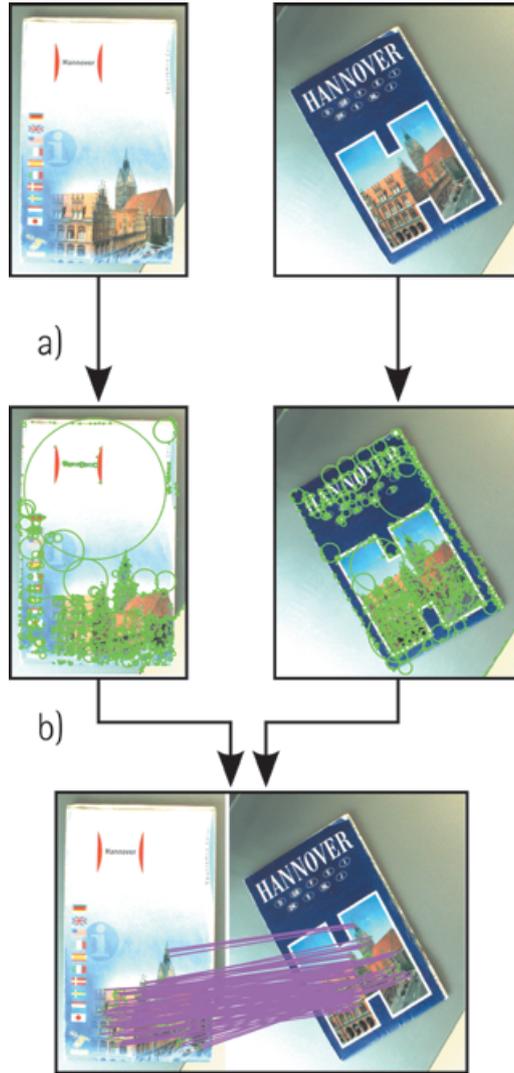


Abbildung 3
Workflow zur Bestimmung von Korrespondenzen zwischen zwei Bildern: **a)** Merkmale (grüne Kreise geben die Ausdehnung an) werden extrahiert. **b)** Übereinstimmende Deskriptoren in korrekter Konstellation sind Korrespondenzen (magentafarbene Linien).
Quelle: Institut für Informationsverarbeitung

Die spannende Frage, die im Projekt vom IMS beantwortet werden soll, ist, welche speziellen Instruktionen für die Videoüberwachung bereitgestellt werden sollten. Dazu muss untersucht werden, wie die Berechnungen bei Extraktion und Vergleich der Bildmerkmale ablaufen. Ebenso kann eine Umorganisation der Abarbeitungsreihenfolge genutzt werden, so dass mehrere Berechnungsschritte zu einer Instruktion zusammengefasst werden können.

Gesellschaftliche Aspekte einer Automatisierung der Situationseinschätzung

Ziel der sozialwissenschaftlichen Untersuchungen am Institut für Soziologie (IfS) ist es, herauszufinden, anhand welcher Kriterien Operateure eines automatischen Systems zur Situationseinschätzung eine Person als verdächtig oder gefährlich klassifizieren. Wie entsteht ein Verdacht, anhand welcher Faktoren werden Personen als potenzielle Risiken eingestuft, welche Akteure sind an dem Prozess der Situationseinschätzung beteiligt und inwiefern werden Personen stigmatisiert oder ausgegrenzt?

Wie bedeutsam eine Analyse der automatischen Situationseinschätzung ist, verdeutlichen sozialwissenschaftliche Studien über die Arbeit von Video-Operateuren. Operateure von nicht-automatisierten Systemen nehmen oftmals Migranten, junge Frauen oder Nicht-Sesshafte grundlos ins Visier. Einzelne gesellschaftliche Gruppen werden aufgrund sozialer oder kultureller Vorurteile überwacht, als »riskant« klassifiziert und ausgegrenzt.

Diese Form des sozialen Aussortierens (social sorting) einzelner gesellschaftlicher Gruppen kann durch eine

bare Instruktionen zerlegt werden. Ein ASIP könnte für die Division eine spezielle Recheneinheit und eine dazugehörige Instruktion bereitstellen. Dies spart Zeit, da für jede Division wesentlich weniger Instruk-

tionen ausgeführt werden müssen, und führt zu einer merklich schnelleren Ausführung des Algorithmus, besonders dann, wenn viele Divisionen in einem Algorithmus vorkommen.

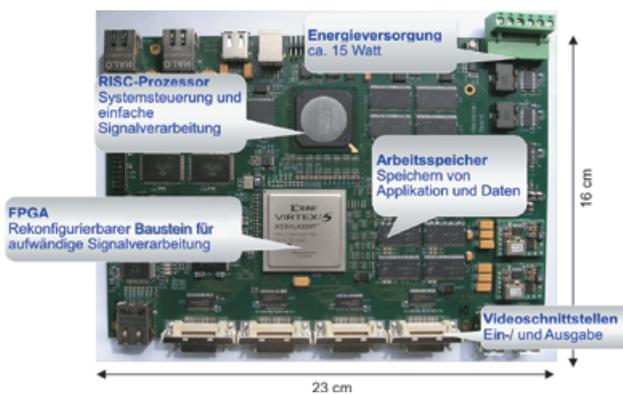


Abbildung 4
Exemplarische Plattform zur digitalen Videosignalverarbeitung
Quelle: Institut für Mikroelektronische Systeme

Automatisierung der Videoüberwachung verstärkt werden. Bei automatisierten Systemen werden die Kriterien der Situationseinschätzung vielfach von den Operateuren in das System eingespeist. Die Einschätzung und Bewertung einer Situation basiert also nicht mehr allein auf konkreter Beobachtung, sondern auf dem automatisierten Abgleich zwischen dem erfassten Merkmalsprofil der beobachteten Person und den in das System eingegebenen Regeln. Vorurteile, Vorannahmen oder subjektive Bewertungen könnten so anonymisiert und automatisiert werden.

Automatische Situationseinschätzung

Das Erkennen und Identifizieren von Objekten oder Personen genügt noch nicht, um bei Gefahrensituationen automatisch Alarm zu schlagen. Die Beziehungen zwischen den Objekten und der Umgebung müssen ebenfalls berücksichtigt werden, zum Beispiel:

»Fährt ein erkanntes Auto auf ein Flugzeug zu?« oder »Befindet sich eine erkannte Person in einem gesperrten Bereich?«. Dabei kann Umgebung sowohl räumlich als auch zeitlich von Bedeutung sein, denn bestimmte Situationen sind nur zu einem bestimmten Zeitpunkt oder an einem bestimmten Ort gefährlich. Das Überwachungssystem muss weiterhin wissen, was genau eine Gefahrensituation ist, denn an verschiedenen Orten gelten verschiedene Situationen als gefährlich. Ein Tankwagen darf neben einem Flugzeug stehen, um das Flugzeug zu betanken, parkt er hingegen auf dem Rollfeld, liegt eine Gefahrensituation vor. In sicherheitskritischen Umgebungen werden diese Zusammenhänge in Form von Sicherheitsregeln abgebildet. Das gilt für Flughäfen, wo diese Regeln vom Überwachungspersonal gekannt und

Abbildung 5
Gruppenbild der Autoren
Hinten: Ralf Dragon, Jörn Ostermann, Holger Blume, Philipp Kärger (von links nach rechts)
Vorne: Carsten Dolar, Matthias Rieger, Fabian Abel (von links nach rechts)



durchgesetzt werden, aber auch für Hochleistungsrechner, wo Sicherheitsregeln von bestimmten Programmen automatisch überprüft werden.

Die Herausforderung der automatischen Situationseinschätzung besteht darin, diese Regeln mit den identifizierten Objekten zu verbinden und anhand von logischen Schlussfolgerungen einzuschätzen, ob eine Gefahr vorliegt. Eine weitere Herausforderung ist die Integration von verschiedenen Datenquellen: Zum Beispiel müssen Flugpläne und Dienstpläne berücksichtigt werden, um zu entscheiden, wann sich bestimmte Personen in der Nähe eines bestimmten Flugzeugs aufhalten dürfen. Am Forschungszentrum L3S wird zurzeit ein solches System entwickelt. Es basiert auf klassischen Techniken der Künstlichen Intelligenz, wie dem regelbasierten Schließen, die mit modernen Verfahren, wie dem sogenannten Stream Reasoning, kombiniert werden, um die zeitliche Komponente beim Schlussfolgern zu berücksichtigen.

Dipl.-Ing. Ralf Dragon

Jahrgang 1980, ist seit 2007 wissenschaftlicher Mitarbeiter am Institut für Informationsverarbeitung und forscht an der computergestützten Analyse von Videomaterial. Neben der Anwendung der Methoden der merkmalsgestützten Bildverarbeitung beschäftigt er sich mit der Erforschung neuartiger Merkmale zur vollständigen Objektbeschreibung. Kontakt: dragon@tnt.uni-hannover.de

Dr.-Ing. Carsten Dolar

Jahrgang 1977, ist seit 2009 wissenschaftlicher Mitarbeiter am Institut für Mikroelektronische Systeme im Fachgebiet »Architekturen und Systeme«. Kontakt: dolar@ims.uni-hannover.de

Prof. Dr.-Ing. Jörn Ostermann

Jahrgang 1962, leitet seit 2003 das Institut für Informationsverarbeitung. Seit 2008 leitet Jörn Ostermann die Requirements Subgroup of MPEG (ISO/IEC JTC1 SC29 WG11). Kontakt: ostermann@tnt.uni-hannover.de

Dr. Matthias Rieger

Jahrgang 1965, ist seit 2008 wissenschaftlicher Mitarbeiter am Institut für Soziologie. Derzeitige Forschungsschwerpunkte sind die sozialen Auswirkungen von Sicherheits- und Überwachungstechnologien sowie die Geschichte und gesellschaftliche Bedeutung der Technisierung der Sinneswahrnehmung. Kontakt: m.rieger@ish.uni-hannover.de

Prof. Dr.-Ing. habil Holger Blume

Jahrgang 1967, leitet seit 2008 das Institut für Mikroelektronische Systeme und ist Professor für das Fachgebiet »Architekturen und Systeme«. Kontakt: blume@ims.uni-hannover.de

M. Sc. Fabian Abel

Jahrgang 1981, arbeitet seit 2006 am Forschungszentrum L3S und beschäftigt sich im Rahmen seiner Promotion mit Benutzermodellierung und Personalisierung im »Social Web«. Kontakt: abel@L3S.de

Dipl.-Inf. Philipp Kärger

Jahrgang 1981, ist seit 2006 Projektleiter und Promotionsstudent am Forschungszentrum L3S. Seine Forschungsschwerpunkte sind die Nutzung semantischer Technologien für die Sicherheit im Internet und Soziale Netzwerke. Kontakt: kaerger@L3S.de