

Seite: 17
Ressort: Wissen
Seitentitel: WISSEN

Ausgabe: Hauptausgabe
Jahrgang: 2016
Nummer: 13

Der fahrbare Weltversther

Selbststeuernde Autos müssen abschätzen können, ob gleich ein Fußgänger über die Straße läuft, oder ob der Vordermann hält oder abbiegt. Dazu muss man dem Bordcomputer praktisches Wissen einbläuen. Ein mühsames Unterfangen. Tübinger Max-Planck-Informatiker haben damit begonnen. Von Christian J. Meier

Webcams gibt es für ein paar Euro. Handys haben mittlerweile zwei Kameras eingebaut. Und Oberklassewagen orientieren sich mit Stereosensoren, die Bilder räumlich erfassen können, ähnlich wie der Mensch. Für Andreas Geiger vom Max-Planck-Institut für Intelligente Systeme in Tübingen bedeuten die allgegenwärtigen Kameras weitaus mehr: Sie sind für ihn die Augen der Computer.

"Wahrnehmung ist ein essenzieller Teil von Intelligenz", sagt der Informatiker. Weil Rechner sich in der Welt des Menschen zurechtfinden sollen - als Haushaltsroboter oder als selbstfahrende Autos - hält Geiger die künstlichen Augen für den wichtigsten Sinn. Doch es gibt ein Problem. Computer verstehen die meisten Bilder nicht. Für sie sind sie nur ein chaotisches Mosaik von Millionen verschiedenfarbiger Pixel und keine Häuser, Bäume, Fußgänger, Autos, Ampeln oder Bordsteine. Zudem schätzen Menschen aufgrund von Bildern Situationen ein. "Davon sind Computer noch weit entfernt", betont Geiger. Um etwa ein Auto ohne Hilfe des Fahrers durch den dichten Stadtverkehr zu lotsen, müssten Rechner beurteilen können, ob der Vordermann gleich abbiegt oder nicht, oder ob das Kind am Straßenrand im nächsten Moment auf die Straße rennt oder nicht. Geiger: "Daher entwickeln wir Systeme, die wie ein Mensch wahrnehmen und entsprechend reagieren können."

Der 33-Jährige und sein vierköpfiges Team arbeiten an einem Programm, mit dem der Computer zweidimensionale Kamerabilder in ein räumliches Modell von der Wirklichkeit umrechnen kann. "Dabei ergeben sich Mehrdeutigkeiten", erklärt Geiger. Wenn sich ein Computer ein Bild erklären soll, auf dem ein dicker Baumstamm zu sehen ist, hat er mehrere Möglichkeiten: Der dicke Stamm könnte nicht nur zu einem Baum gehören, sondern sich zusammensetzen aus zwei dünneren Stämmen, die leicht

schräg versetzt hintereinander stehen. Weil dem Kamerabild die räumliche Tiefe fehlt, kann die künstliche Intelligenz das nicht sicher unterscheiden.

Mehrdeutigkeit entsteht für den Computer beispielsweise auch in einer von Altbauten gesäumten Wohnstraße, die auf beiden Seiten mit Autos zugeparkt ist, wie Geiger anhand von Fotos demonstriert. Die Bilder sind mit leicht unterschiedlichem Winkel aufgenommen; so ähnlich arbeiten die Augen des Menschen, damit das Gehirn daraus einen räumlichen Eindruck ableiten kann.

Ähnlich schätzt ein Computer Entfernungen ein, indem er etwa misst, wie weit verschoben ein Fensterrahmen ist, wenn er die zwei Aufnahmen von der Straßenszene miteinander vergleicht. Normalerweise wertet der Rechner dabei direkt die Daten aus: Er sucht für jedes Pixel auf dem einen Bild den zugehörigen Partner im anderen. Um den Fensterrahmen hier wie dort aufzuspüren, vergleicht der Computer die Farbwerte der einzelnen Bildpunkte. "Kanten wie ein Fensterrahmen lassen sich auf diese Weise leicht orten", meint Geiger. Sie zeigen einen abrupten Übergang von einer Farbe zur anderen, der sich überall leicht wiedererkennen lässt. Der Lack an der Autotür dagegen ist meistens einfarbig; alle Pixel haben einen ähnlichen Farbwert. Vor dieser Mehrdeutigkeit kapituliert die Standard-Software: Sie kann die Autos nicht auseinanderhalten.

Deshalb geben die Tübinger Forscher ihrem Bilderklärungsprogramm Wissen mit über das, was es da sieht: Sie machen aus einer Ansammlung von Bildpunkten eine Szene. Zum Beispiel werden für den Computer jene Bereiche im Bild markiert, wo sich Autos befinden. Dann simuliert die Software die Wagen dreidimensional und reiht sie hintereinander auf. So lässt sich eine durchgängige Tiefenkarte erzeugen. Die Entfernungen und Räume müssen nicht mit unterschiedlichen Kamerablickwin-

keln abgeschätzt werden.

Aber ganz klar ist die Sache damit immer noch nicht. Denn das Programm erkennt nicht, wie viele Autos am Straßenrand stehen, welche Fahrzeugtypen es sind und wie eng oder parallel sie zur Bordsteinkante parken. Es gibt also Tausende von Variationen, die man in Tübingen durchtestet. Geiger: "Auf diese Weise wird die wahrscheinlichste Hypothese herausgefiltert."

Indem die Forscher Erfahrung und Gesehenes zusammenführen, lernt der Computer, die Welt zu deuten. "Dabei ist es wichtig, dass man das Problem insgesamt betrachtet und nicht nur seine einzelnen Bestandteile", sagt Geiger. Um Gegenstände und Personen auf einem Bild richtig in Beziehung zu setzen, bringt man dem Rechner außerdem bei, dass sich Autos und Fußgänger nicht durchdringen. Erst dieses Weltwissen ermöglicht es dem Computer, bewegte Bilder sinnvoll einzuordnen. Geiger nennt das den "Szenenfluss". Der ist für den Computer nur schwer nachzuvollziehen, denn, so Geiger: "Aus 1 Meter 60 Höhe, wo die Stereokameras am Auto angebracht sind, lässt sich dieses Wissen nicht ableiten." Oft sehe die Kamera im Wagen nicht einmal, ob die Ampel gerade auf Rot oder Grün steht. Damit das Auto solche Szenen versteht, bräuchte man eine Kamera, die das Ganze aus der Vogelperspektive filmt. Dann würden sich für den Computer nur die Fahrzeuge bewegen, er wüsste, wer auf welcher Spur fährt, welche Ampeln es gibt und wie sie gerade geschaltet sind.

Doch diese Draufsicht hat der Wagen nicht, also müssen die Tübinger Experten den Bordcomputer schlauer machen. Wie bringt man ihm zum Beispiel bei, was ein Fußgänger und was ein fahrendes oder parkendes Auto ist? Für den Rechner, der über die Stereokameras sieht, bewegt sich nämlich alles, weil sich seine eigenen Kameras mitbewegen, wenn der Wagen fährt. Man kann

dem Computer die Regel eingeben, dass ein Auto nicht mit den Armen schlenkert, sondern sich als kompaktes Ganzes bewegt. Schon ist es für den Rechner einfacher, die Szene zu verstehen: Er kann viele falsche Annahmen streichen, etwa die, dass ein Auto durch einen Laternenmast zweigeteilt und dann zu zwei Objekten wird.

Um der Software diese Art von praktischer Intelligenz mitzugeben, setzen die Wissenschaftler auf eine Methode namens "Maschinelles Lernen". Dabei üben Computer anhand von vielen Bildern, bestimmte Elemente zu erkennen. Das macht man sich schon in der Gesichtserkennung zunutze: Das Programm wird trainiert. Es bekommt Tausende Porträtfotos präsentiert und kann am Ende auf neuen Fotos selbstständig bestimmte Gesichter ausmachen.

Ganz ähnlich lernt Geigers Software, einzelne Wagen auf einer Kreuzung zu entdecken und sie zu verfolgen: Fahren sie geradeaus? Biegen sie ab? Aus den Erfahrungswerten, wie sich die anderen Autos verhalten, kann der Computer auf den Verkehrsfluss schließen und von da auf Ampeln, Fahrspuren oder Abbiegestreifen. Auch das Umfeld der Kreuzung wird gescannt: Wo stehen

Gebäude, wo sind die Fluchtpunkte? Mit all diesen Informationen bastelt sich der Rechner eine digitale Karte von der Kreuzung zusammen und lässt vor seinem inneren Auge einen 3-D-Film ablaufen, der die von den Kameras eingefangene Szenerie erklärt.

"Durch die Kombination von Kameras und Intelligenz kämen autonome Fahrzeuge ohne die aufwendige Technik aus, die heutige Prototypen mit sich führen, etwa Laserscanner oder Radar", meint Geiger. Auch Satellitennavigation und genaue digitale Karten seien nicht mehr nötig. Und da ohnehin nicht damit zu rechnen sei, dass überall am Straßenrand für viel Geld aufwendige Technik installiert wird, die autonome Autos unterstützt - zumindest so lange nicht, wie es kaum selbstfahrende Wagen gibt -, hätte der Tübinger Do-it-yourself-Ansatz viel Charme. "Wir wollen aufwendige Technik durch künstliche Intelligenz ersetzen", sagt Geiger.

Nun ja, aber da ist auch noch der freie Bürger, der gern freie Fahrt für sich reklamiert und der es besser kann als alle anderen. Ob er sich künftig von seinem Wagen chauffieren lassen will? "Akzeptanz", nennt Geiger das. Die werde kommen, "sobald die Systeme

deutlich weniger Fehler machen als ein Mensch", glaubt der Informatiker. Im Moment allerdings sei der Computer noch lange nicht so gut wie ein Autofahrer.

Selbstbild Das Fernziel der Wissenschaftler ist, dass die Wagen künftig selber mitdenken im Straßenverkehr - ganz ohne Satelliten und teure Installationen am Straßenrand. (foto: imago)

