

3

Wie ein Roboter funktioniert



Was in diesem Kapitel behandelt wird

Dieses Kapitel wurde für Leser geschrieben, die keinen großen technischen Hintergrund zu intelligenter interaktiver Robotik haben.

- Die grundlegenden Hardware- und Softwarekomponenten, aus denen ein Roboter besteht.
- Die Techniken, die wir anwenden können, um einen Roboter für die Interaktion mit Menschen bereit zu machen.

Wie funktioniert ein Roboter? Lassen sie uns dafür ein Rollenspiel spielen: Wir stellen uns vor, ein Roboter zu sein. Zunächst denken wir vielleicht, wir können viele Dinge tun, aber bald stellen wir fest, dass unsere Fähigkeiten stark eingeschränkt sind. Wenn wir ein neu gebauter Roboter sind, ohne entsprechende Software, sind unsere Gehirne völlig leer. Wir können nichts tun – weder uns bewegen, noch wissen, wo wir sind, noch verstehen, was um uns herum ist, noch um Hilfe bitten. Wir empfinden diese Erfahrung als ziemlich seltsam und schwer vorstellbar. Die Hauptquelle der Seltsamkeit ist, dass das Gehirn des neuen Roboters nicht mit dem eines Menschen vergleichbar ist, nicht einmal mit dem eines Säuglings. Der Roboter hat keine grundlegenden Instinkte, keine Ziele, kein Gedächtnis, keine Bedürfnisse, keine Lernfähigkeiten und keine Fähigkeit, zu fühlen oder zu handeln. Für die Herstellung eines Robotersystems müssen wir Hardware und Software integrieren und zumindest teilweise gemeinsam entwickeln, damit der Roboter fühlen und handeln kann.

In diesem Kapitel betrachten wir die üblichen Komponenten eines Roboters und wie sie miteinander verbunden sind, um die Teilnahme an einer Interaktion zu ermöglichen. In Abschnitt 3.1 werden grundlegende Ideen zu den Komponenten erläutert, die zum Bau eines Roboters benötigt werden. In Abschnitt 3.2 werden verschiedene Arten von Hardware erläutert. Abschnitt 3.3 stellt Sensoren wie Kameras, Entfernungsmesser und Mikrofone, Abschnitt 3.4 stellt Aktoren vor. Abschnitt 3.5 schließlich erklärt die Software, die diejenigen Hardware-Elemente begleitet, die sich mit der Wahrnehmung (z.B. Computer-Sehen), Planung und Handlungssteuerung des Roboters befasst.

■ 3.1 Die Herstellung eines Roboters

Für den Bau eines Roboters besteht einer der ersten Schritte darin, Verbindungen zwischen den Sensoren, dem Computer und den Motoren des Roboters herzustellen, sodass der Roboter in der Lage ist, zu fühlen, zu interpretieren, was er fühlt, Aktionen zu planen und diese dann auszuführen. Sobald der Roboter z. B. mit einer Kamera verbunden ist, kann sein Computer die von der Kamera gelieferten Daten lesen. Das Kamerabild ist jedoch nichts anderes als eine große Zahlentabelle, ähnlich der folgenden Tabelle:

9	15	10
89	76	81
25	34	29

Können Sie anhand dieser Zahlen erraten, was der Roboter sieht? Vielleicht einen Ball, einen Apfel oder eine Gabel? Wenn man davon ausgeht, dass jeder Wert in der Tabelle den Helligkeitswert eines Sensorelements in der Kamera darstellt, können wir diese Zahlen in eine Grafik übersetzen, die für den Menschen aussagekräftiger wird (siehe Bild 3.1), für den Roboter aber bedeutungslos bleibt.

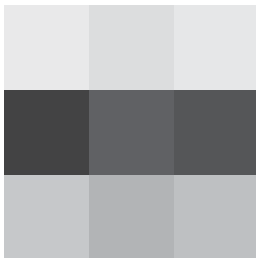


Bild 3.1

Die Daten der Kamera werden in ein Raster aus Graustufenpixeln übersetzt.

Sie können vielleicht eine Linie in Bild 3.1 sehen, aber ein Roboter hat kein Verständnis dafür, was eine Linie ist. Diese Linie könnte die Kante einer Klippe sein, von der der Roboter fallen und sich selbst beschädigen könnte. Der Roboter hat jedoch kein Konzept der Höhe oder der Schwerkraft. Er würde nicht begreifen, dass er fallen könnte, wenn er diese Linie überquert. Er weiß nicht, dass er im Falle eines Sturzes wahrscheinlich auf dem Kopf stehen würde. Er würde nicht einmal erkennen, dass sein Arm gebrochen wäre. Mit anderen Worten, selbst Konzepte, die für die Interaktion mit der Welt um uns herum und das Überleben in dieser Welt von entscheidender Bedeutung sind und die dem Menschen angeboren sind, müssen in einen Roboter explizit einprogrammiert werden.

Ein Roboter ist im Wesentlichen ein Computer mit einem Körper. Jede Funktionalität muss in den Roboter programmiert werden. Ein Problem, mit dem alle Roboter

zu tun haben, ist, dass ihre Sensoren und Motoren zwar für den Betrieb in dieser Welt ausreichen, ihre Intelligenz jedoch nicht. Jedes Konzept, das für Robotik-Forscher von Interesse ist, muss verinnerlicht, d. h. in den Roboter programmiert werden. Dieses erfordert viel Zeit und Mühe und beinhaltet oft viele Zyklen aus Versuch und Irrtum. Die analoge Welt der Wirklichkeit wird in eine digitale Welt umgewandelt, und die Übersetzung von Zahlentabellen in aussagekräftige Informationen und sinnvolle Antworten ist eines der Kernziele der künstlichen Intelligenz. Die Fähigkeit, ein Gesicht aus einer großen Wertetabelle zu identifizieren, zu erkennen, ob eine Person schon einmal gesehen wurde, und den Namen dieser Person zu kennen, sind alles Fähigkeiten, die programmiert oder gelernt werden müssen. Daher wird der Fortschritt der Mensch-Roboter-Interaktion (HRI) durch die Fortschritte im Bereich der künstlichen Intelligenz eingeschränkt. Robotik-Ingenieure integrieren Sensoren, Software und Aktoren, um dem Roboter zu ermöglichen, seine physische und soziale Umgebung zu verstehen und mit ihr zu interagieren. Ein Ingenieur könnte zum Beispiel Beschleunigungssensoren verwenden, die die Beschleunigung und die Erdanziehung erfassen können, z. B. um die Ausrichtung des Roboters zu lesen und festzustellen, ob er gefallen ist. Ein Klippensensor, der aus einer kleinen, nach unten gerichteten Infrarot-Lichtquelle und einem Lichtsensor besteht, kann vom Roboter verwendet werden, um einen Sturz von einer Treppe zu vermeiden.

Zu den typischen Problemen, die Robotik-Ingenieure für den Roboter lösen müssen, gehören die folgenden:

- Was für einen Körper hat der Roboter? Hat er Räder? Verfügt er über Arme?
- Wie erkennt der Roboter seine Position im Raum?
- Wie steuert und positioniert der Roboter seine Körperteile – z. B. Arme, Beine, Räder?
- Wie sieht der Raum um den Roboter herum aus? Gibt es Hindernisse, Klippen, Türen? Was braucht der Roboter, um diese Umgebung wahrnehmen zu können und sich sicher in ihr zu bewegen?
- Was sind die Ziele des Roboters? Woher weiß er, wann er sie erreicht hat?
- Sind Menschen in der Nähe? Wenn ja, wo sind sie, und wer sind sie? Wie kann der Roboter das wissen?
- Schaut eine Person den Roboter an? Spricht jemand mit ihm? Wenn ja, was versteht der Roboter durch diese Hinweise?
- Was versucht der Mensch zu tun? Was will der Mensch, dass der Roboter es ausführt? Wie können wir sicherstellen, dass der Roboter dies versteht?
- Was soll der Roboter tun und wie soll der Roboter reagieren?

Um diese Fragen anzugehen, müssen HRI-Forscher geeignete Hardware und eine geeignete Morphologie für den Roboter bauen oder auswählen und dann entspre-

chende Programme – die Software – entwickeln, die dem Roboter sagen können, was er mit seinem Körper tun soll.

■ 3.2 Roboter-Hardware

Als dieses Buches geschrieben wurde, sind bereits eine Reihe von Robotern für den Verbrauchermarkt produziert worden. Auch wenn nicht alle von ihnen zu Haushaltsgeräten geworden sind, eignen sich diese kommerziellen Roboter oft als Plattformen für die HRI-Forschung. Kommerziell erhältliche Roboter bieten eine Vielzahl von Körpertypen, darunter tierähnliche, humanoide und eher mechanische.

Aibo, ein Beispiel für einen tierähnlichen Roboter, sieht aus wie ein Hund mit einem etwas mechanischen Aussehen (siehe Bild 3.2). Er kann sehen, hören, fühlen, berühren, Geräusche machen, mit den Ohren und dem Schwanz wedeln und sich auf seinen vier Beinen bewegen. Die ersten Aibo-Modelle wurden 1999 verkauft; 2006 wurde der Verkauf eingestellt. Seit 2017 sind neue Modelle lieferbar.



Bild 3.2

Roboter Aibo ERS-1000
(2018 – heute). (Quelle: Sony)

Pepper hingegen ist ein humanoider Roboter in der Größe eines Heranwachsenden (siehe Bild 3.3). Einige Geschäfte verwenden Pepper, um Besucher anzuziehen und Waren und Dienstleistungen zu vermarkten. Das Unternehmen, das Pepper herstellt, hat auch den kleineren humanoiden Roboter Nao (siehe Bild 2.4) zum Kauf durch Endverbraucher im Angebot.

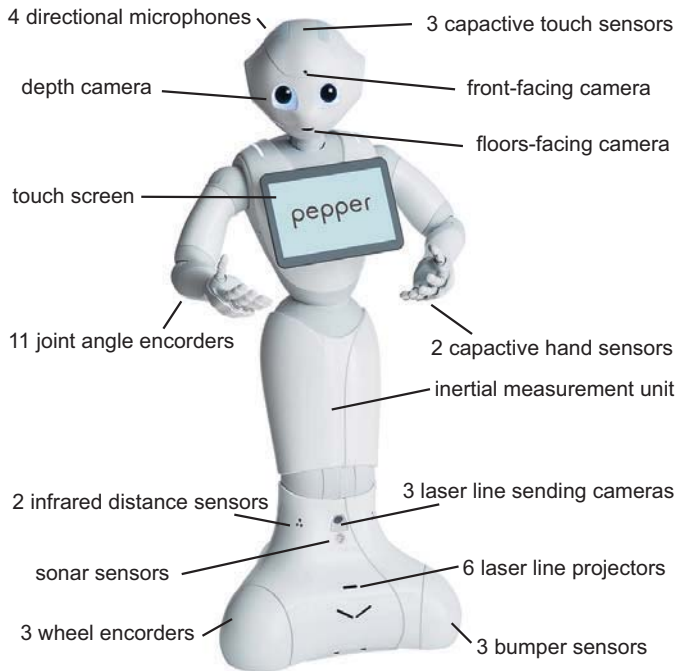


Bild 3.3 Pepper Roboter (2014 - heute) und seine Sensoren. (Quelle: Softbank Robotics und Philippe Dureiltoma)

Der eher mechanisch aussehende Sicherheitsroboter K5 ist in den Vereinigten Staaten im Handel erhältlich und einer der wenigen Roboter, die für den Einsatz im Freien bestimmt sind.

Roboter, die nicht explizit für den Einsatz in der HRI konzipiert wurden, können dennoch für HRI-Studien eingesetzt oder sogar modifiziert werden. Der kommerziell erfolgreichste Heimroboter ist nach wie vor der Staubsaugerroboter iRobot Roomba (siehe Bild 9.1), der weltweit millionenfach verkauft wurde. Roombas sind nicht nur ein interessantes Mittel, um die Beziehung der Öffentlichkeit zu Robotern zu untersuchen (Forlizzi und DiSalvo, 2006), sondern wurden auch für die HRI-Forschung modifiziert und umprogrammiert. iRobot stellt eine programmierbare Version des Roomba her, den Create, dem die Staubsaugerkomponente fehlt und der in der Forschung und im Bildungsbereich für Roboter eingesetzt wird.

Telepräsenzroboter können ebenfalls als Plattformen für die HRI-Forschung eingesetzt werden. Es gibt viele verschiedene Typen auf dem Markt, darunter mobile Versionen wie den Beam und Desktop-Versionen wie Kubi. Derzeit werden kleine

mobile Roboter mit einem Bildschirm zur Marktreife entwickelt, die ein freundliches Gesicht zeigen.

Obwohl kommerziell erhältliche Roboterhardware eine Vielzahl von Morphologien sowie sensorische und programmierbare Fähigkeiten bietet, ist jeder Roboter in seinen Möglichkeiten begrenzt; sein Aussehen und seine Fähigkeiten schränken die Interaktionen ein, die er durchführen kann. Forscher konzipieren und bauen daher auch ihre eigenen Roboter, die von einfachen Desktop- und mobilen Plattformen mit oder ohne Manipulator bis hin zu sehr menschenähnlichen Androiden reichen. Die Wahl einer bestimmten Morphologie für einen Roboter, der in der HRI-Forschung eingesetzt werden soll, hängt oft von den Fähigkeiten ab, die für die erwartete Aufgabe benötigt werden (z. B. ob er in der Lage sein soll, Objekte aufzunehmen), von der Art der Interaktion (z. B. können tierähnliche Interaktionen von einem tierähnlichen Roboter profitieren) und von den Erwartungen und Wahrnehmungen der Menschen in Bezug auf verschiedene Morphologien (z. B. kann von Humanoiden erwartet werden, dass sie sich ähnlich wie Menschen verhalten und ähnlich intelligent sind).

■ 3.3 Sensoren

Die meisten sozialen Roboter sind mit Sensoren ausgestattet, die es ihnen ermöglichen, das Geschehen in ihrer Umgebung zu erfassen. Viele der üblicherweise verwendeten Sensoren beziehen sich auf die drei am häufigsten verwendeten Modalitäten in der menschlichen Interaktion: Sehen, Hören und Berühren. Aber Roboter sind keineswegs auf menschliche Wahrnehmungsarten beschränkt. Es ist daher oft hilfreich, zu überlegen, welche Arten von Informationen der Roboter wahrnehmen muss und welche die genauesten und zweckmäßigsten Möglichkeiten sind, damit er dies tun kann, anstatt sich auf die Reproduktion menschlicher Fähigkeiten zu konzentrieren.

3.3.1 Sehfähigkeit

Kamera

Eine Kamera besteht aus Linsen, die ein Bild auf eine Sensorfläche fokussieren. Die Sensoroberfläche wird entweder mit einem ladungsgekoppelten Bauelement (CCD) oder – häufiger – mit einer komplementären Metalloxid-Halbleiter-Technologie (CMOS) verwirklicht. Das Grundelement einer Kamera ist ein Lichtsensor, der hauptsächlich aus Silizium besteht und Licht in elektrische Energie umwandelt.

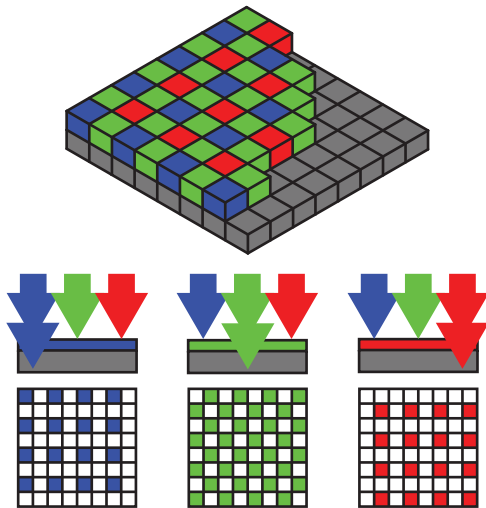


Bild 3.4
Eine Reihe von CCDs in einer RGB-Kamera.

Eine Kamera besteht aus Millionen dieser Lichtsensoren. Normalerweise wird die Farbe in einem Kamerabild durch drei Werte dargestellt, nämlich rot (R), grün (G) und blau (B). Daher wird eine Kamera allgemein als RGB-Kamera bezeichnet. Die Sensoren auf der Sensoroberfläche sind nicht empfindlich für die Farbe des auf sie auftreffenden Lichts; sie sind nur für die Lichtintensität empfindlich. Für den Bau einer RGB-Kamera werden kleine Farbfilter auf der Sensoroberfläche angebracht, wobei jeder Filter nur rotes, grünes oder blaues Licht durchlässt (siehe Bild 3.4). Kameras sind die mächtigsten und komplexesten Sensoren, die Robotern zur Verfügung stehen. Ihre weite Verbreitung in Digitalkameras und Smartphones hat dazu geführt, dass RGB-Kameras sehr klein und billig geworden sind.



In der Forschung zum Computer-Sehen setzen Forscher häufig Kameras in der räumlichen Umgebung ein, um eine präzise Sehfähigkeit zu ermöglichen. Obwohl dies einer der realistischen Ansätze für die Erreichung einer stabilen Leistung des Computer-Sehens ist, wird davon in der HRI-Umgebung manchmal abgeraten, weil sich Menschen in der Nähe von sichtbaren Kameras unwohl fühlen können. Bei einem Projekt, in dem ältere Menschen zu Hause von einem Roboter unterstützt werden sollten, hätten die Ingenieure beispielsweise gerne Kameras am Roboter und in der Wohnung gehabt, weil der Roboter dann die Menschen genau verfolgen und mit ihnen interagieren könnte. Die Teilnehmer lehnten jedoch die Installation und den Einsatz von Kameras entschieden ab, was das Team zwang, stattdessen Lokalisierungssignale und Laserentfernungsmesser zu verwenden (Cavallo et al., 2014).

Die meisten Kameras haben ein eingeschränkteres Sichtfeld als der Mensch. Während Menschen mehr als 180 Grad sehen können, sieht eine typische Kamera etwa nur 90 Grad und verpasst so viel von dem, was sich in der Peripherie abspielt. Ein

Roboter mit einer einzigen Kamera hat ein begrenztes Sichtfeld und muss sich möglicherweise auf andere Sensoren wie Laserentfernungsmesser oder Mikrofone verlassen, um einen Eindruck dafür zu bekommen, was um ihn herum passiert.

Wichtig ist, dass das Kamerabild erst mithilfe von Algorithmen für das Computer-Sehen verarbeitet werden muss, damit der Roboter auf seine visuelle Umgebung reagieren kann (siehe Abschnitt 3.5.4).

Tiefensensoren

Genauso wie menschliches Sehen in Stereo funktioniert und das Wissen über gesehene Objekte sowie die Eigenbewegung nutzt, um die Entfernung zu den anvisierten Objekten zu ermitteln, können Algorithmen für das Computer-Sehen verwendet werden, um ein dreidimensionales (3D-) Bild aus zweidimensionalen (2D) Informationen zu extrahieren. Stereokameras waren lange Zeit die Technologie der Wahl, aber in den letzten Jahren sind Technologien entstanden, die es uns ermöglichen, Tiefe direkt zu sehen, ohne dass wir auf das Computer-Sehen angewiesen sind. Diese „Tiefensensoren“ erzeugen ein „Tiefenbild“ bzw. RGBD-Bild (wobei D für Tiefe, engl. depth, steht), das eine Karte von Entfernungen zu Objekten im Blickfeld der Kamera darstellt.

Normalerweise kann ein Tiefensensor den Abstand zu Objekten messen, die wenige Meter entfernt sind. Abhängig von der Stärke des emittierten Infrarotlichts funktionieren die meisten Tiefensensoren nur in Innenräumen zuverlässig. Es gibt verschiedene Möglichkeiten für die Herstellung solcher Tiefensensoren. Eines der typischen Verfahren ist das Laufzeitverfahren (Time of Flight, TOF), bei dem ein Gerät unsichtbare Impulse aussendet und die Zeit zwischen dem Moment, in dem es das Licht aussendet, und dem Moment, in dem es die Reflexion des Lichts empfängt, misst. Die aktuelle Elektronik-Hardware ist aufgrund der hohen Lichtgeschwindigkeit nicht imstande, die Zeit des normalen Lichts mit ausreichender Präzision aufzeichnen, stattdessen sendet die Kamera Infrarotlichtimpulse aus und misst dabei die Phasendifferenz. Microsoft Kinect One, die zweite Iteration von Microsofts Game Controller, basiert auf diesem Prinzip (siehe Bild 3.5). Obwohl er ursprünglich als Spiele-Controller entwickelt wurde, ist er schnell von den Roboterbauern benutzt worden und wird heute weithin verwendet, um Robotern einen Eindruck von Tiefe zu vermitteln. In Kombination mit entsprechender Software kann der Kinect-Sensor auch eine Skelettverfolgung durchführen, was hilfreich ist, um herauszufinden, wo sich Menschen befinden, was sie tun und sogar, wie sie sich fühlen. Es sind jetzt kleinere Geräte erhältlich, die RGBD-Bilder auf der Grundlage verschiedener Technologien, einschließlich TOF, strukturiertem Licht und Stereosehen, zurückgeben.



Bild 3.5 Microsoft Kinect Azure DK Sensor für Windows. (Quelle: Microsoft)

Laser-Entfernungsmesser

Für die Messung von Entfernungen in größeren Dimensionen verwenden Forscher häufig einen Laser-Entfernungsmesser, der auch als Light Detection and Ranging (LIDAR) bezeichnet wird. Ein typischer Laser-Entfernungsmesser kann Abstände zu Objekten messen, die bis zu 30 Meter entfernt sind. Er tastet dabei die Umgebung zwischen 10 und 50 Mal pro Sekunde ab. Die Genauigkeit der Auflösung von Laser-Entfernungsmessern liegt bei wenigen Zentimetern. Der grundlegende Mechanismus dieses Sensortyps ist ebenfalls TOF. Ein Laser-Entfernungsmesser sendet einen einzelnen Strahl infraroten Laserlichts aus und misst die Entfernung, indem er die Zeit zwischen dem Moment des Sendens des Laserstrahls und dem Zeitpunkt, an dem er seine Reflexion empfängt, misst. Typischerweise befinden sich Sender und Empfänger auf einer rotierenden Plattform, mit der der Laserstrahl durch die Umgebung geführt wird. Daher misst das Gerät die Entfernung nur in einer einzigen 2D-Ebene (d. h. in der Rotationsebene der rotierenden Plattform).

Bei Robotern können Entfernungsmesser in verschiedenen Höhen montiert sein, um in einer horizontalen Ebene nach Objekten zu suchen. Entfernungsmesser in Bodennähe können Objekte am Boden und an den Beinen von Personen erfassen, während höher angebrachte Entfernungsmesser zur Erfassung von Objekten auf einem Tisch oder einer Ablagefläche verwendet werden können (siehe Bild 3.6).

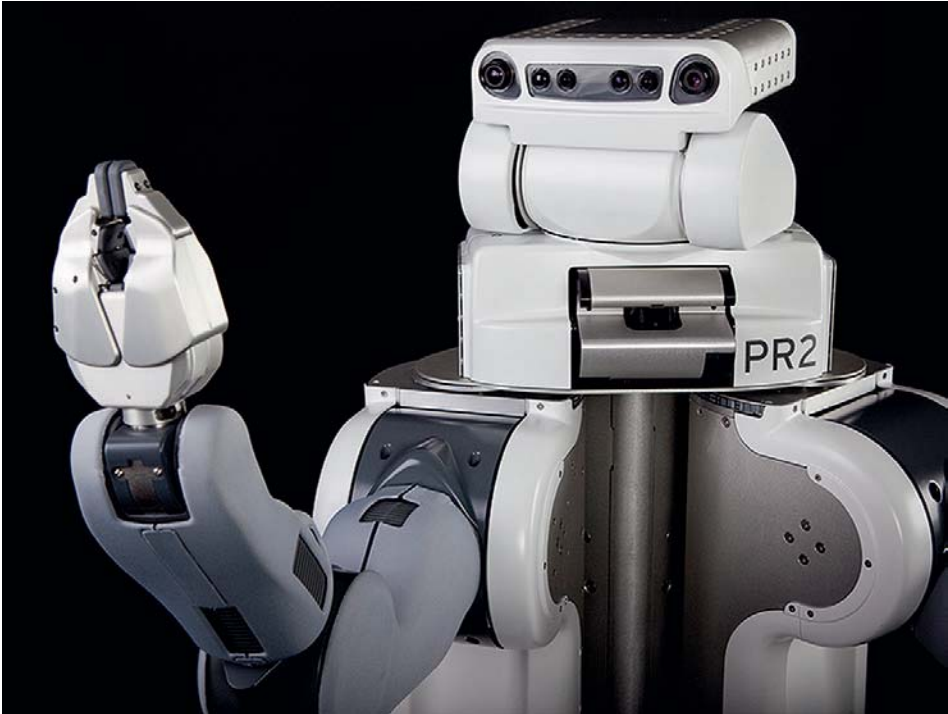


Bild 3.6 Die PR2-Roboter (2010 – 2014): Können Sie sagen, wo der Entfernungsmesser ist? (Quelle: Willow Garage)

3.3.2 Audio

Für die akustische Wahrnehmung werden üblicherweise Mikrofone verwendet. Sie wandeln Schall in elektrische Signale um. Mikrofone haben unterschiedliche Empfindlichkeitsprofile; einige sind omnidirektional, d. h. sie nehmen alle Geräusche in der Umgebung auf, während andere gerichtet sind und nur in einem kegelförmigen Bereich vor dem Mikrofon Geräusche aufnehmen. Die Kombination mehrerer Mikrofone ermöglicht die Anwendung von „Strahlformungs“-Techniken, sodass Schallsignale aus einer bestimmten Richtung von Umgebungsgeräuschen unterschieden werden können. Mikrofone in Reihen werden zur Schallquellenlokalisierung verwendet, d. h. zur genauen Bestimmung des Winkels einer Schallquelle in Bezug auf deren Position im Verhältnis zu den Mikrofonen.

3.3.3 Berührungssensoren

Berührungssensoren, bzw. taktile Sensoren können in der HRI wichtig sein, wenn ein Roboter physisch vom Nutzer geführt werden soll. Es gibt viele verschiedene Implementierungen, von physischen Tasten oder Schaltern bis hin zu kapazitiven Sensoren, wie sie auf Touchscreens zu finden sind (siehe Bild 3.7).

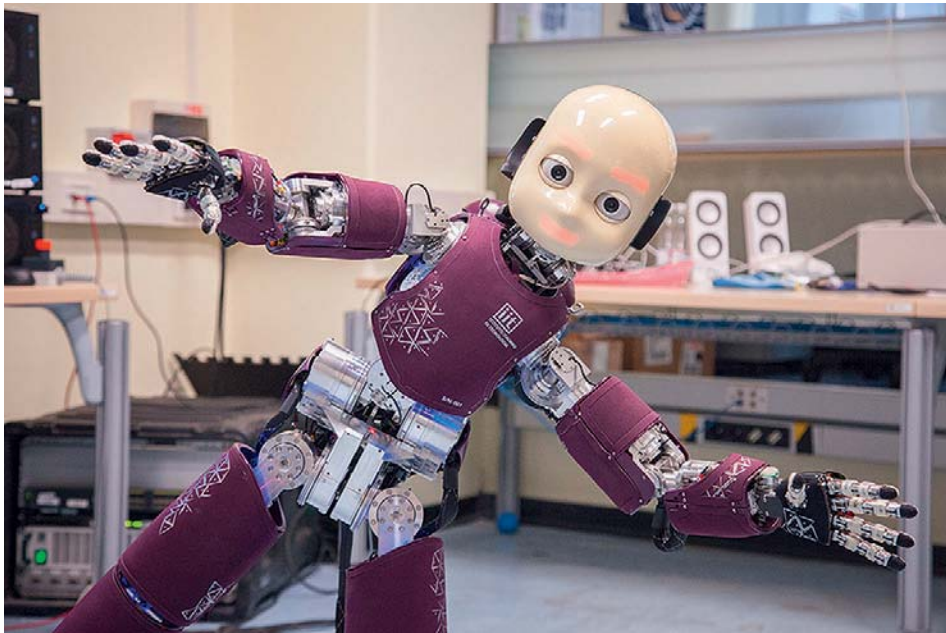


Bild 3.7 Der iCub Humanoid (seit 2004) hat kapazitive taktile Sensoren in den Fingern, Handflächen und im Rumpf eingebaut. (Quelle: Dynamic Interaction Control)

Der am häufigsten verwendete taktile Sensor ist ein mechanischer Druckschalter. Er wird oft zusammen mit einem Stoßfänger verwendet. Wenn ein Roboter mit einem Objekt kollidiert, wird der Schalter geschlossen, sodass der Roboter die Kollision erkennen kann. Drucksensoren und Kapazitätssensoren, wie diejenigen, die die Position Ihres Fingers auf einem Berührungsbildschirm ablesen, können auch zur Erkennung von physischem Kontakt mit der Umgebung verwendet werden. Drucksensoren können mithilfe verschiedener Technologien implementiert werden, enthalten aber normalerweise ein Material, das seine elektrischen Eigenschaften (Widerstand oder Kapazität) bei Anwendung von Kraft ändert. Drucksensoren können Robotern helfen zu erkennen, ob und wie stark sie eine Person oder einen Gegenstand berühren. Sie sind auch sehr nützlich, um Roboter in die Lage zu versetzen, Objekte angemessen aufzunehmen und zu handhaben. Taktile Sensoren können auch verwendet werden, damit der Roboter erkennt, ob jemand ihn