

Dem Geräusch auf der Spur

WIE ROBOTER HÖREN, WO ETWAS PASSIERT

Das menschliche Gehör ist ein Meisterwerk an Präzision. Diese Fähigkeiten auf einen Roboter zu übertragen, ist eine wissenschaftliche Herausforderung. Forscher vom Institut für Informationsverarbeitung stellen drei Methoden vor, mit denen der Roboter Geräusche lokalisieren können soll, um dem Sinneseindruck »Hören« näher zu kommen.

Das Ziel bei der Entwicklung von Robotern ist es, dass diese in irgendeiner Weise mit der Umwelt interagieren. Hierzu braucht der Roboter so etwas wie Sinneseindrücke. Das heißt, der Roboter soll beispielsweise sehen oder hören können. Um Robotern Sinneseindrücke zu verleihen, werden Sensoren benötigt, die Messwerte liefern. Mikrofone sind solche Sensoren und ein erster Baustein, damit Roboter hören können. Die Messwerte der Mikrofone sind allerdings ohne eine Interpretation nutzlos. Durch die digitale Aufnahme eines Geräusches entstehen zwar mehrere zehntausend Werte pro Sekunde an Daten zum Schalldruck – sie sind aber für den »hörenden« Roboter zunächst nicht zu gebrauchen. Eine nützliche Information müsste zum Beispiel darüber Auskunft geben, welches Geräusch aufgetreten oder an welchem Ort das Geräusch entstanden ist – erst dann können Entscheidungen getroffen oder sinnvolle Handlungen eingeleitet werden. Diese erwünschten Informationen sind in den Zahlenkolonnen der Messwerte bereits vorhanden, allerdings versteckt. Mit einer Verarbeitung der Signale können sie zum Vorschein gebracht werden. Eine solche Signalverarbeitung ist damit sehr wichtig, um einem Roboter das Hören beizubringen.

Der erste Schritt, bevor ein Geräusch lokalisiert werden

Lokalisation und Ortung

Lokalisation und Ortung sind zwei Begriffe, die fälschlicherweise oft synonym verwendet werden. Beide Begriffe beschreiben, wie die Position eines Objektes bestimmt wird. Der Unterschied liegt in dem Verfahren der Ortsbestimmung sowie darin, von welchen Objekten der Ort bestimmt werden kann.

Der Begriff **Ortung** wird verwendet, wenn der Beobachter oder die Beobachterin für das Ermitteln des Ortes eines Objekts aktiv Wellen sendet und gleichzeitig Wellen empfängt. Die ausgesendeten Wellen müssen dabei von dem Objekt reflektiert werden. Anhand dieser zurückkommenden Wellen wird die Position des Objekts ermittelt. Die Ortung funktioniert also nur, wenn das Objekt in geeigneter Weise die ausgesendeten Wellen reflektiert.

Der Begriff **Lokalisation** wird verwendet, wenn für das Ermitteln des Ortes eines Objekts der Beobachter oder die Beobachterin ankommende Wellen empfängt, aber zuvor keine Wellen aktiv gesendet hat. Wenn das Objekt selbst Wellen abstrahlt, werden diese vom Beobachter empfangen und mit diesen Informationen wird der Ort des Objekts bestimmt. Bei der Lokalisation kann also nur der Ort eines Objekts bestimmt werden, wenn von diesem Wellen ausgehen. Menschen lokalisieren mit ihren Augen und Ohren.

Dieser Artikel behandelt die Lokalisation mittels Schallwelle. Der gesonderte Themenkomplex der Ortung wird hier nicht betrachtet.

kann, ist das automatische Erkennen des Geräusches. Danach kann bestimmt werden, woher das Geräusch kommt. Dies ist für viele Aufgaben wichtig, zum Beispiel bei der Überwachung von Objekten auf Schäden. Hier ist die Lokalisation entscheidend, um den Schaden richtig einzuordnen. Das Institut für Informationsverarbeitung entwickelt erste Ansätze für die Lokalisation von Schäden an Rotorblättern von Windener-

gieanlagen mittels Luftschall. Dies geschieht im Rahmen eines Projekts zur Schadensfrüherkennung zusammen mit dem Institut für Statik und Dynamik und Prof. Dr.-Ing. habil. Raimund Rolfes, welcher das Projekt leitet. Ein Mensch mit zwei gesunden Ohren lokalisiert ständig Geräusche, beispielsweise ein hupendes Auto oder ein klingelndes Mobiltelefon. Roboter werden immer häufiger in einer für den Menschen ge-

schaffenem Umgebung eingesetzt, sollen mit Menschen interagieren und sich wie ein Mensch verhalten. Ein Roboter, der sich im Straßenverkehr autonom bewegt, sollte daher erkennen, woher Geräusche kommen. Auch ein humanoider, also menschenähnlicher Serviceroboter benötigt die Fähigkeit, zu lokalisieren. Wenn dem Roboter beispielsweise eine Frage aus einer Menschengruppe heraus ge-

die Lokalisation im dreidimensionalen Raum werden mindestens vier Mikrofone benötigt, um den Ort der Quelle zu bestimmen. Sind weniger Mikrofone vorhanden, kann der Entstehungsort lediglich eingegrenzt werden. In dem Kasten »Lokalisation mittels Laufzeitunterschieden« ist das Konzept veranschaulicht. Das Lokalisationskonzept funktioniert sehr gut, wenn es keine Nebengerä-

Mit dem Konzept **Beamforming** wird der Schall aus jeder Richtung einzeln betrachtet. Dazu wird ein Aufbau von vielen Mikrofonen verwendet, welcher Mikrofonarray genannt wird. Das Konzept der Signalverarbeitung ist hier, alle Mikrofonensignale in bestimmter Weise zu verzögern und zu überlagern. Beamforming wird häufig für so genannte akustische Kameras verwendet. Hier wird Schall in einem

Abbildung 1–6
Lokalisation mittels Laufzeitunterschieden: Die Lokalisation ist in der Realität eine Aufgabe, welche im dreidimensionalen Raum gelöst werden muss. Zur besseren Anschauung ist hier die Lokalisation in einer zweidimensionalen Ebene gezeigt.

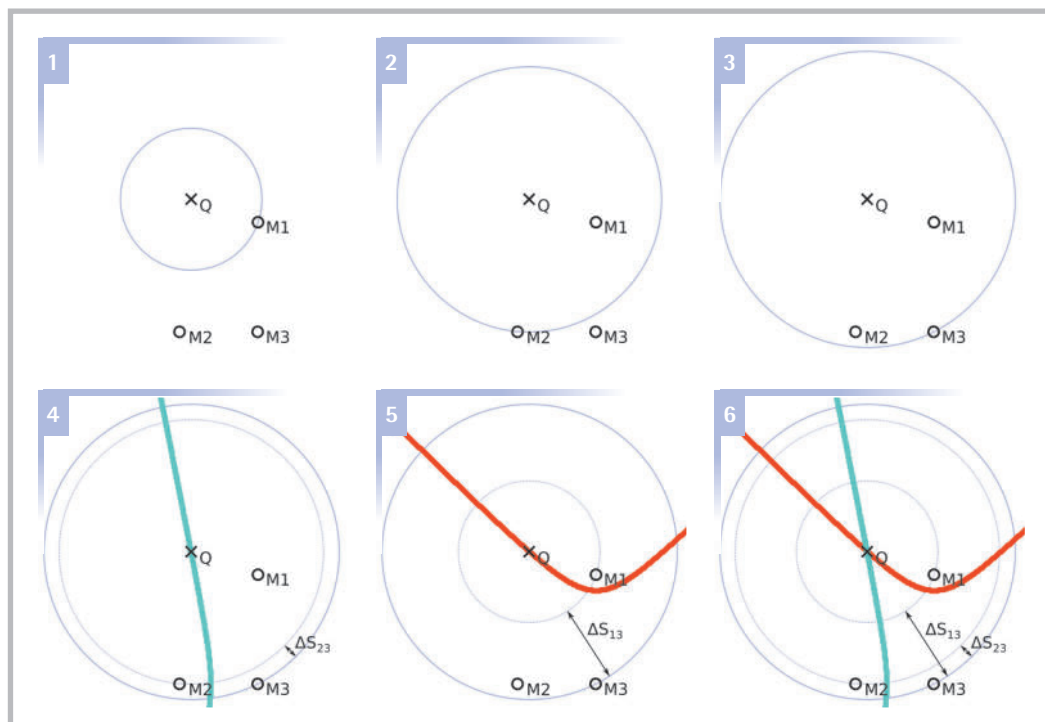


Abbildung 1
Die Welle, welche von der Quelle Q ausgesendet wurde, kommt zum Zeitpunkt Z1 an dem Mikrofon M1 an.

Abbildung 2
Zum Zeitpunkt Z2 erreicht die Welle das Mikrofon M2.

Abbildung 3
An dem Mikrofon M3 kommt die Welle zum Zeitpunkt Z3 an.

Abbildung 4
Aus den Ankommenszeitpunkten Z2 und Z3 berechnet man die Streckendifferenz ΔS_{23} . Mit dieser Differenz lässt sich ein Hyperbelzweig (blau) berechnen, auf welchem die Quelle liegen muss.

Abbildung 5
Mit den Ankommenszeitpunkten Z1 und Z3 lässt sich die Streckendifferenz ΔS_{13} berechnen und damit ein zweiter Hyperbelzweig (rot) bestimmen. Auch auf dieser Kurve muss die Quelle liegen.

Abbildung 6
Der Schnittpunkt der beiden Hyperbelzweige ist das Ergebnis des Lokalisationsverfahrens. Anhand der Grafik ist zu sehen, dass die Position der Quelle richtig bestimmt wurde.

stellt wird, sollte der Roboter sich der richtigen Person zuwenden.

Es gibt drei wichtige Methoden, mit denen Geräusche lokalisiert werden können: Lokalisation mittels Laufzeitunterschieden, Beamforming sowie Lokalisation durch Pegelunterschiede. Für alle Verfahren werden mehrere Mikrofone benötigt. Das wichtigste Konzept ist die **Lokalisation mittels Laufzeitunterschieden**. Ein Laufzeitunterschied entsteht dadurch, dass die Schallwelle eine unterschiedliche Strecke zu den Mikrofonen zurücklegt. Für

sche und keine Reflexionen der Schallwellen gibt. Reflexionen und Nebengeräusche sind aber bei allen realen Einsatzbedingungen vorhanden und erschweren bei allen Verfahren die Lokalisation. Des Weiteren kann sich sowohl das Objekt als auch der Roboter bewegen. Auch dies muss beachtet werden und führt zu komplexeren Verfahren. Eine noch größere Herausforderung entsteht, wenn keine Sichtlinie zwischen dem Geräusch und dem Mikrofon besteht. Am Institut für Informationsverarbeitung wird unter anderem an Lösungen für diesen Fall geforscht.

Bildausschnitt zusammen mit dem optischen Bild dargestellt. So kann beispielsweise erkannt werden, an welcher Stelle ein Motor Lärm erzeugt. Im Gegensatz zu der Lokalisation mittels Laufzeitunterschieden können mit Hilfe des Beamformings nur die Richtungen, aus welchen der Schall auf das Mikrofonarray getroffen ist, bestimmt werden. Die Entfernung der Quelle ist so nicht zu ermitteln. Dafür ist der Vorteil des Beamformings das Handhaben von mehreren Quellen aus unterschiedlichen Richtungen, die zur gleichen Zeit Schall abstrahlen.



Prof. Dr.-Ing. Jörn Ostermann

Jahrgang 1962, ist seit 2003 Leiter des Instituts für Informationsverarbeitung. Seine Forschungsinteressen sind Audio- und Videosignalverarbeitung, Computer Vision, 3D Modellierung, Gesichtsanimation und Mensch-Maschine-Interaktion. Kontakt: ostermann@tnt.uni-hannover.de



Thomas Krause

Jahrgang 1985, ist seit 2013 Wissenschaftlicher Mitarbeiter am Institut für Informationsverarbeitung. Seine Forschungsinteressen sind die Audio-Signalverarbeitung, Mustererkennung und Lokalisation und die automatisierte Zustandsüberwachung. Kontakt: krause@tnt.uni-hannover.de

hierbei die Eigenschaften der Ausbreitung von Schallwellen beachtet werden, um richtige Schlüsse zu ziehen. Dieser Mechanismus ist zum Beispiel für das Lokalisieren einer Schallquelle im Nahbereich nützlich, da etwa unterhalb eines Meters oft relevante Pegelunterschiede auftreten.

Ein Teilbereich der Robotik verfolgt das Ziel, Roboter zu erschaffen, welche so menschenähnlich wie möglich sind. Die menschliche Fähigkeit Geräusche zu lokalisieren, ist äußerst komplex – daher ist die menschliche Signalverarbeitung sehr schwer zu imitieren. Ein Mensch verfügt über lediglich zwei Trommelfelle, welche den Schall ähnlich wie zwei Mikrofone aufnehmen. Damit ist es mit nur einem Lokalisationskonzept nicht

Mensch eine beeindruckende räumliche Lokalisationsschärfe auf. Ein wichtiger Effekt, den Menschen hierfür ausnutzen, basiert darauf, dass sehr viele Geräusche gleichzeitig unterschiedliche Tonhöhen enthalten. Durch die Form und Anordnung der beiden Ohren werden Geräusche, die von oben, von vorne oder jeder anderen Richtung kommen, verändert und zwar in Abhängigkeit der Tonhöhe. Bei einer festen Richtung ist die Veränderung immer gleich. Diese Information wird für die Lokalisation verwendet. Die Auswertung allein dieses einen Effekts ist schon sehr komplex und zeigt beeindruckend, wie gut der Mensch Signale verarbeitet. Eine ähnliche Lokalisationsgenauigkeit mit zwei Ohrsignalen künstlich zu erreichen ist noch in weiter Ferne und umfasst zahlreiche Herausforderungen für die Signalverarbeitung.

Zusammenfassend kann festgehalten werden, dass die Lokalisation von Geräuschen für Roboter ein wichtiger Teil des »Sinneseindrucks« Hören darstellt. Es gibt schon ausgereifte Lokalisationsverfahren, welche in bestimmten, meist einfachen oder klar beschränkten Einsatzbereichen sehr gut funktionieren. Für sehr viele Roboteranwendungen ist die Vielfalt der auftretenden Bedingungen aber riesig und es ergeben sich häufig Bedingungen, welche schwierig zu handhaben sind. Zudem können sich die Bedingungen bei mobilen Robotern auch noch schnell ändern. Die Lokalisation muss aber idealerweise in allen Szenarien robust funktionieren. Mit der Forschungsarbeit der Audiolokalisation werden Lösungen erdacht, welche die bisherigen Hürden überwinden. Diese Lösungen tragen dazu bei, zukünftige Robotergenerationen zu erschaffen, die besser hören und deshalb neue Aufgaben übernehmen können.

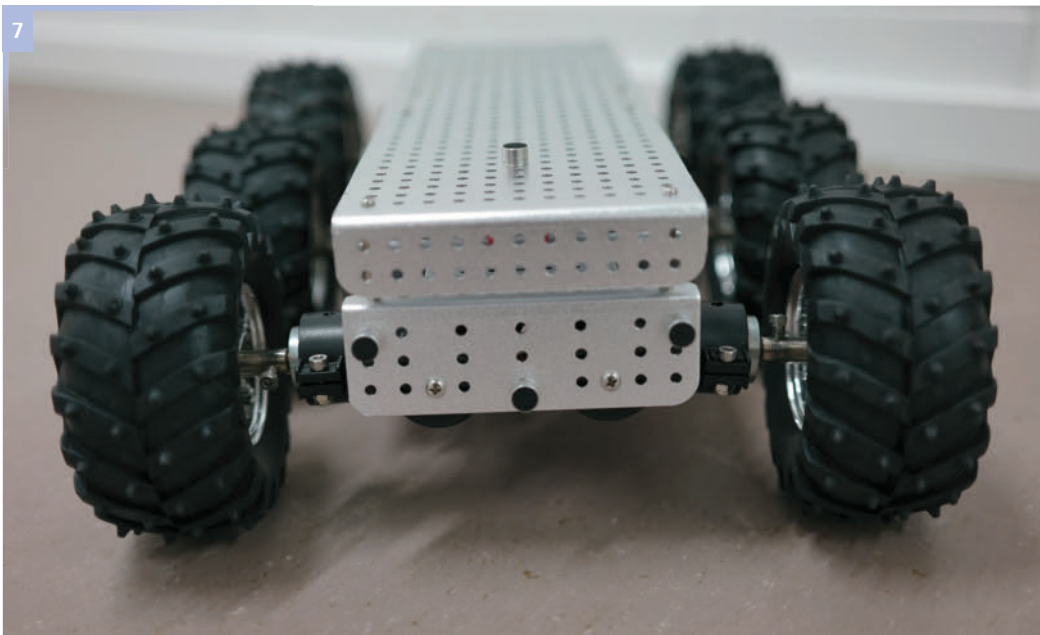


Abbildung 7
Roboterplattform mit drei Mikrofonsensoren an der Front und einer Kapsel auf dem Dach

Ein weiteres Verfahren ist die **Lokalisation mittels Pegelunterschieden**. Hier werden Schalldrücke, die von mehreren Mikrofonen aufgefangen werden, verglichen. Anhand dieser Daten wird der Ort der Schallquelle ermittelt. Dabei ist ein höherer Pegel ein Indiz für einen geringeren Abstand zur Quelle. Allerdings müssen

möglich, Geräusche räumlich zu lokalisieren. Dies kann anschaulich in dem Kasten »Lokalisation mittels Laufzeitunterschieden« nachvollzogen werden. Der Mensch verwendet daher viele unterschiedliche Lokalisationskonzepte gleichzeitig, um Geräusche räumlich zu lokalisieren. Für nur zwei Ohren weist der