

WILLST DU WOHL HÖREN?

*EIN INTERNATIONALES FORSCHERTEAM
UNTER DER LEITUNG DER TU BERLIN WILL ROBOTERN
MENSCHLICHES HÖREN BEIBRINGEN.*

*Text: Oliver Geyer
Foto: www.picture-alliance.com
Illustration: Jochen Michael*

ABER DAS IST EIN SEHR KOMPLEXES PHÄNOMEN, FÜR DAS ES EINER GUTEN VERNETZUNG BEDARF. IN DER TECHNIK WIE AUCH IM TEAM.

— Nachdem sie als blinkende Blechbüchsen fast schon auf dem Schrotthaufen ausgedienter Zukunftsvisionen gelandet waren, sind sie heute wieder voll da: die Roboter. Sie sind intelligenter geworden – und vernetzter. In Südkoreas Hauptstadt Seoul gibt es Baustellen, auf denen hunderte Roboter in enger Kooperation beim Errichten ganzer Hochhäuser helfen. Robotik-Experten haben ihnen beigebracht, nach dem Vorbild von Ameisenstaaten, via Sensoren und Funkchips so miteinander zu kommunizieren, dass sie Aufgaben bewältigen, für die jede einzelne Maschine zu dumm ist. Doch was Robotern bis heute nicht so recht gelingen will, ist eine menschliche Wahrnehmung. Das EU-Forschungskonsortium Two!Ears arbeitet nun daran, das Gehör des Menschen mithilfe von Robotern nachzuahmen. Vernetzter kann ein Projekt kaum sein: Neun Forschungsinstitute aus ganz Europa und den USA arbeiten zusammen, um die auditive Wahrnehmung des Menschen zu simulieren – die ihrerseits mit dem gesamten Sinnesapparat verwoben ist. Darüber haben wir mit Professor Alexander Raake von der Technischen Universität Berlin gesprochen, der das Projekt leitet.

JEDE MENSCHLICHE WAHRNEHMUNG FUNKTIONIERT JA ÜBER EIN NETZWERK VON NERVENZELLEN. GEHT ES BEI DEM PROJEKT TWO!EARS DARUM, DIESES NETZWERK FÜR DAS MENSCHLICHE GEHÖR NACHZUBAUEN?

— Das neuronale Netz eins zu eins physiologisch nachzubauen, ist nicht unser Ziel, das können wir nicht leisten. Aber wir forschen an einem Testsystem, mit dem sich das menschliche Hören technisch funktional nachbilden lässt.

WAS GENAU BEABSICHTIGEN SIE DAMIT?

— Erstmal ist das Grundlagenforschung, was wir hier betreiben. Es sind also die verschiedensten Einsatzmöglichkeiten denkbar. Von uns entwickelte Algorithmen könnten mal in Hörgeräten genutzt werden. Auch ein Einsatz in der Robotik selbst ist vorstellbar, etwa bei Rettungsrobotern, die in einer natürlichen Umgebung Opfer orten sollen. Neben Instrumenten wie Ultraschall und Infrarot könnte dabei auch eine menschliche Form der auditiven Wahrnehmung hilfreich sein. Kann ein Roboter in einer komplexen Schallsituation erfassen, dass da jemand nach Hilfe ruft? In der Interaktion mit Menschen ist eine menschliche Hörfähigkeit sicher hilfreich. Noch anspruchsvoller ist unser Ziel, die Mechanismen nachzubilden, mit denen Menschen beim Hören Qualität wahrnehmen. Dabei ist die Robotik erstmal nur Mittel zum Zweck, um das menschliche Gehör nachzubilden, das in Vernetzung mit dem gesamten Wahrnehmungsapparat funktioniert.

WIE MUSS MAN SICH

DIESE VERNETZUNG VORSTELLEN?

— Es gibt den so genannten Cocktailparty-Effekt: Sitzt jemand in einer Umgebung mit sehr vielen Geräuschquellen, kann er sich durch seine auditive Aufmerksamkeit auf ganz bestimmte Schallquellen in dieser Umgebung aktiv ausrichten. Zum einen natürlich durch eine Kopfdrehung in Richtung der Schallquelle. Aber er tut dazu auch Dinge, die von außen gar nicht sichtbar sind. Denn unser Gehör hängt also nicht nur von den Schallwellen ab, die auf die Ohren treffen, sondern von vielen anderen Funktionalitäten, die miteinander verwoben sind: davon, welche Bedeutung unser Gehirn ihnen zuschreibt, von unserer taktilen und sensorimotorischen Wahrnehmung, von unserem »Gesichtssinn«, abhängig von der Stellung der Augen und von unserer Körper- und Kopfposition. Neurophysiologisch weiß man, dass es eine direkte Kopplung zwischen der so genannten Sensorimotor-Steuerung des Kopfes und unserer Signalerkennung beim Hören gibt. Man nimmt an, dass es Kombinationen aus bestimmten Schallmustern und Kopfpositionen gibt, die im Hirn als Referenzen abgelegt sind und eine besonders effiziente Ortung und Erkennung von Geräuschen ermöglichen. Solche entsprechenden Top-Down-Prozesse zur Steuerung der Bewegungen oder der Signalanalyse müssen wir für eine Simulation des menschlichen Gehörs einbeziehen.

UND WIE SIMULIEREN SIE EIN

SO KOMPLEXES PHÄNOMEN TECHNISCH?

— Da müssen wir mit ganz verschiedenen Mitteln arbeiten. Die Drehung des Kopfes in Richtung der Schallquelle etwa, die der Mensch für die Exploration des Schallfeldes braucht, stellen wir mit einem Roboter dar – mit einem so genannten Kopf-und-Torso-Simulator. Das ist ein anthropomorphisch weitgehend korrekter Kopf, der auf einem Stock nachgebildeten Torso angebracht ist und zu Rotationen in der Lage ist. Das Taktile, den Gleichgewichtssinn und das Gefühl für die eigene Position hat so ein Roboter sozusagen built-in. Für andere die Funktionalitäten des Hörens bauen wir ein komplexes kognitives Softwaresystem, das zum Beispiel die Konzentration des Menschen auf den eigenen Namen bestimmte Stimmen oder Geräusche simuliert. Wobei man sagen muss: Eine Gehör-Expertise, wie sie der Mensch im Laufe seiner Kindheit entwickelt, werden wir so einer Maschine nie beibringen können. Die visuellen Daten schließlich, die für das Hören auch eine Rolle spielen, erfassen wir über ein Kamerasystem. Wir schaffen also mit technischen Mitteln ein vernetztes Gebilde, das diese verschiedenen Funktionalitäten integriert und dadurch dem menschlichen Gehör ähnlich wird. Bistlang hat man in der Robotik zwar auch schon eine recht gut umfassende Signalverarbeitung eingesetzt, die aber meist mit dem Installieren einer Vielzahl von Mikrofonen umgesetzt wurde. Das ist etwas anderes als das menschliche Gehör mit lediglich zwei Ohren, die aber wie richtungsbasierte Filter wirken.

**AUCH WENN EIN ROBOTER BEIM HÖREN NICHT
SO SCHNELL DIE FÄHIGKEITEN EINES MENSCHEN
ERWERBEN WIRD, WIE BRINGEN SIE IHM
ÜBERHAUPT BEI, MENSCHENÄHNLICH ZU HÖREN?**

— Wir haben Leute von der Ruhr-Uni Bochum sowie der TU Berlin im Team, die sich mit maschinellem Lernen beschäftigen: Die entwickeln ein Expertensystem, das tatsächlich ansatzweise gewisse Ähnlichkeiten mit der neuronalen Struktur des Menschen hat und das trainiert werden kann. Zwei andere Institute aus unserem Konsortium – eins von der TU DTU Dänemark Berlin in Kopenhagen und eins von der Universität in Sheffield – extrahieren Merkmale und realisieren eine computergestützte auditorische Szenenanalyse. Da geht es darum, mehrere Geräuschquellen in einem Schallfeld einzeln auszuwerten und voneinander zu trennen. Dieses Wissen wird nun wiederum losgelassen auf das Substrat, das die Leute von dem maschinellen Lernen bereit stellen. In Verbindung beider Ansätze wird so eine Art Datenbanksystem entwickelt, das einem Spracherkennung ähnelt und das auf verschiedenen Kognitionsstufen über Expertise verfügt. Zum Beispiel über die Fähigkeit der Mustererkennung: Aha, das ist eine Stimme und die ruft jetzt gerade Hilfe. Das System würde verstehen, dass es sich auf Stimmerkennung einstellen muss, obendrein würde es erkennen, aus welcher Richtung der Hilfeschrei kam, und es würde die Geräuschquelle orten. Dieser Prozess entwickelt sich sukzessive.

**DAS HÖREN IST ALSO EIN PHÄNOMEN GUTER VERNETZUNG,
DAS SIE ALS KONSORTIUM VON NEUN WISSENSCHAFT-
LICHEN INSTITUTEN MIT MEHR ALS 20 MITARBEITERN
ERFORSCHEN. WIE VERNETZEN SIE SICH MITEINANDER,
UM DAS ZU BEWERKSTELLIGEN?**

— Wir haben wöchentliche Telefonkonferenzen, die immer unterstützt werden durch eine vernetzte Wirklichkeit mit Texten und Bildern im internen Wiki und in Googledocs, die alle sehen können. Bei diesen Telefonkonferenzen geht es viel um Organisatorisches und die Arbeiten des Projekts. Für jedes Arbeitspaket gibt es dann einen zuständigen Leiter, der zusammen mit den anderen Arbeitsgruppenleitern die Gesamtleitung des Projektes bildet – eigentlich eine gängige Projektmanagement-Struktur. In bestimmten Abständen müssen wir Berichte über unsere Arbeiten an die EU schicken. Das sind zugleich Meilensteine, bei denen wir auch der Presse und der Forschungsgemeinschaft unsere Ergebnisse bereit stellen. Dies geschieht über unsere Website www.twoears.eu, die wir auch als Plattform für unseren internen Austausch nutzen. Und dann beinhaltet das Projekt natürlich wahnsinnig viel Software-Entwicklung. Da arbeiten wir unter anderem mit dem sehr fortschrittlichen Versionsmanagement Git, über das alle Beteiligten auf ein Werkstück im Netz zugreifen können. Das System bietet die Möglichkeit, dass jeder seine eigene Version kreiert, dann aber nur einen Teil dieser Version in den Hauptteil zurückspielt. Nicht zu vergessen: Alle drei Monate treffen wir uns an einem Ort und arbeiten für mehrere Tage direkt zusammen. Das ist ja auch immer noch eine sehr effektive Form der Vernetzung.

**KANN MAN SAGEN, DASS JEDES FORSCHUNGSINSTITUT
EINE DER FUNKTIONALITÄTEN EINBRINGT, DIE IN
IHRER VERNETZUNG DAS GENIALE MENSCHLICHE GEHÖR
ERGEBEN?**

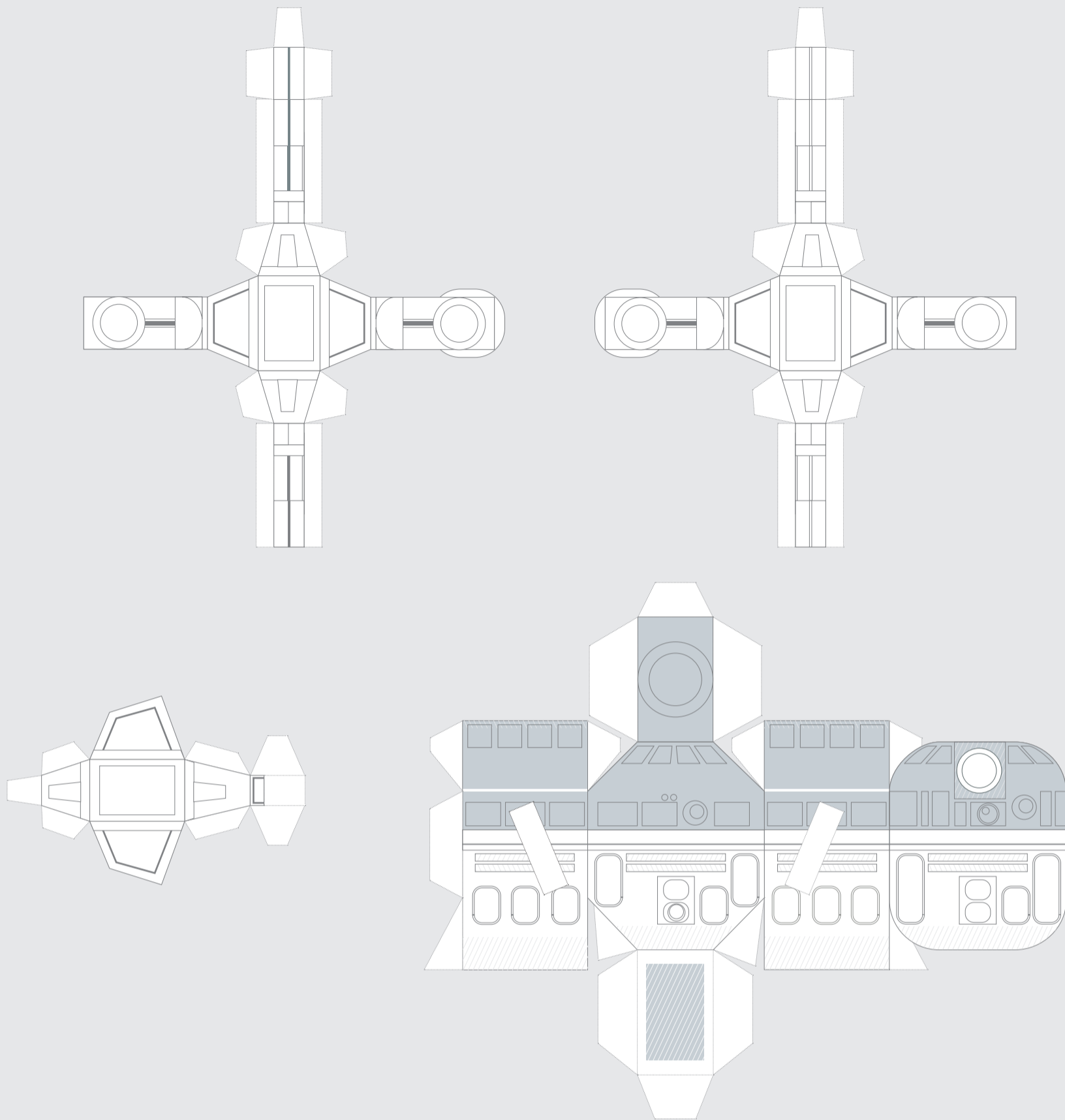
— Das ist ein sehr schönes Bild, das im Grunde auch so zutrifft. Genau genommen sind es aber meist ein bis zwei Partner pro Funktionalität, nicht um Redundanzen zu erzeugen, sondern weil die Aufgabe so komplex ist. So wie das Gehör.

1. ASIMOVSCHE ROBOTERGESETZ

**EIN ROBOTER DARF KEINEN MENSCHEN
VERLETZEN ODER DURCH UNTÄTIGKEIT ZU SCHADEN
KOMMEN LASSEN.**

2. ASIMOVSCHE ROBOTERGESETZ

**EIN ROBOTER MUSS DEN BEFEHLEN
DER MENSCHEN GEHORCHEN, ES SEI DENN,
SOLCHE BEFEHLE STEHEN IM WIDERSPRUCH
ZUM ERSTEN GESETZ.**



3. ASIMOVSCHEs ROBOTERGESETZ

***EIN ROBOTER MUSS SEINE EIGENE EXISTENZ
SCHÜTZEN, SOLANGE DIESES NICHT DEM ERSTEN
ODER ZWEITEN GESETZ WIDERSPRICHT.***

