

Sprachsteuerung für die Industrie

Produktivität und Sicherheit auf Zuruf steigern

In industriellen und medizinischen Anwendungen führen Sprachsteuerungssysteme aktuell noch ein Schattendasein, obwohl sie Anwendern gerade dort das Leben erleichtern können. Wie sich auch Cybersecurity-gerechte Sprachsteuerungssysteme aufbauen lassen, erklären Dr. Diane Hirschfeld, CEO, und Martin Wenzel, Kaufmännischer Leiter von voice Inter connect.

Markt&Technik: Welche Rolle wird Sprachsteuerung im Mix der Bedienverfahren in Anwendungen wie Industrie, Medizin, Gebäudeautomatisierung und Verkehr künftig spielen?

Martin Wenzel: Siri, Alexa und Co. haben zu einem regelrechten Boom von Sprachsteuerungen in Consumer-Produkten geführt. Smartphones und Entertainment-Systeme werden heutzutage ebenso intuitiv per Sprache bedient wie Smart-Home-Anwendungen oder Telefon und Navigationssystem im Automobil. Sprachassistenten sind also längst in unserem Alltag angekommen. Das ist auch kaum ver-

wunderlich: Sprache ist für uns Menschen die natürlichste Form, sich auszudrücken. Das macht sich auch die Industrie verschiedener Branchen zunutze und integriert Sprachbedienung in ihre Lösungen – vom Operationssaal, wo der Operateur so freie Hände behält, über die ablenkungsfreie Bedienung von Fahrzeugen bis hin zu Geräten und Anlagen, bei denen Sprachbefehle komplexe und verschachtelte Funktionen auf Zuruf erreichbar machen. Dabei müssen für Industrieanwendungen deren besondere Anforderungen berücksichtigt werden – vor allem in Bezug auf Zuverlässigkeit, Safety und Security. Lösungen, die eine dau-

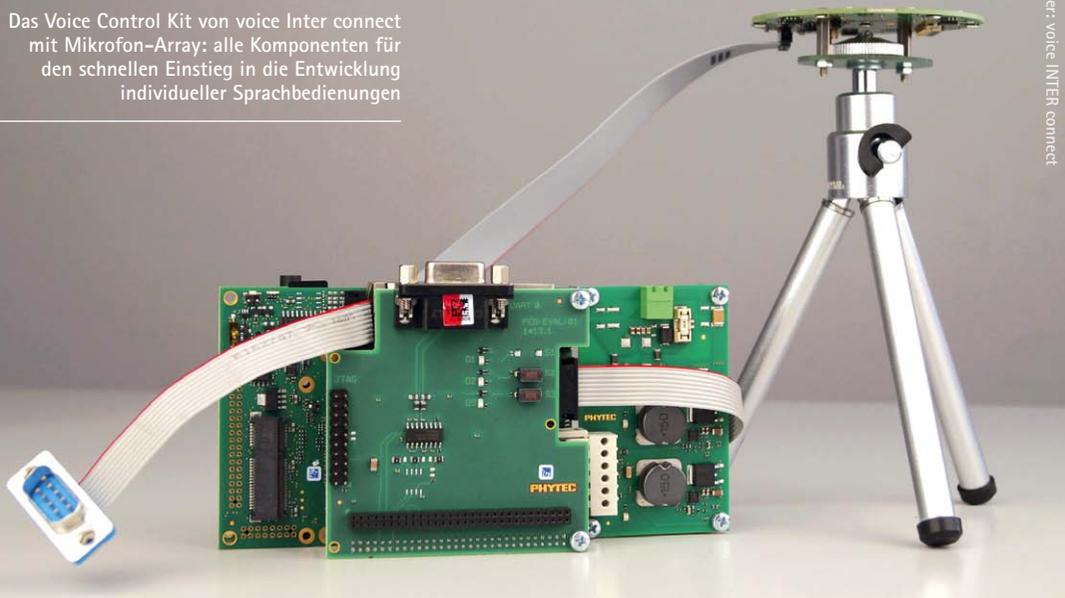
erhafte Cloud-Anbindung benötigen, kommen dafür meist nicht in Frage.

Was kann Sprachsteuerung in solchen Anwendungen derzeit und in absehbarer Zukunft leisten?

Wenzel: Zunächst einmal ist Sprachsteuerung für industrielle Anwendungen einfach, komfortabel und intuitiv. Stellen Sie sich ein zwanzigstöckiges Bürogebäude mit zehn Unternehmen pro Etage vor, in dem Sie einen Termin zum Vorstellungsgespräch haben und nicht wissen, auf welcher Etage das Unternehmen sitzt. Wäre es nicht toll, Sie könnten einfach nur sagen, wohin Sie wollen, und der Fahrstuhl hält in der richtigen Etage? Das funktioniert mit Sprachsteuerung optimal. Dabei evaluiert eine Natural-Language-Understanding/Processing-Einheit Bedeutung und Kontext erkannter Worte und übersetzt sie in entsprechende Steuersignale. Aus der Aussage »Ich möchte bitte gern das komplette Licht anschalten« wird so beispielsweise die Steuerungsabsicht »Licht, 100 Prozent« ermittelt. So lassen sich aus Sprachbefehlen erkannte Werte wie Maschinenparameter, Zeitangaben und Zahlenwerte an Steuerungseinheiten weitergeben.

Selbstverständlich ist es auch mit industrietauglichen Lösungen möglich, für den Start einer Sprachbedienung entsprechende Signalfelder zu programmieren, als Alternative zu einem Tastendruck. Darüber hinaus arbeiten wir an intelligenten Sensoren, die Bestandteil von Maschinen und Gebäuden und sensibel für

Das Voice Control Kit von voice Inter connect mit Mikrofon-Array: alle Komponenten für den schnellen Einstieg in die Entwicklung individueller Sprachbedienungen



Bilder: voice INTER connect

die Anwesenheit von Menschen sind. Sie ermöglichen es auf Basis des Situations- und Bedienkontexts, menschliche Sprachbefehle noch besser von Hintergrundgeräuschen und zwischenmenschlicher Kommunikation zu trennen. Nach der Spracheingabe erkennt ein Dialogmanager, in dem die Logik von Steuerungsvorgängen hinterlegt ist, ob Informationen fehlen oder Befehle nicht valide sind. In dem Fall kann ein akustisches Signal, eine Sprachausgabe oder eine Rückmeldung über die Steuerung weitere Informationen erfragen.

Die Leistungsfähigkeit von Offline-Lösungen wie „vicControl industrial“ ist dabei Lösungen auf Cloud-Basis fast ebenbürtig. In beiden Fällen kommen Deep Neural Networks als Machine-Learning-Methode zum Einsatz. Dabei soll mithilfe künstlicher neuronaler Netzwerke (KNN) die Funktionsweise des menschlichen Gehirns nachgebildet werden. Ziel hierbei ist es, ein selbstlernendes System zu schaffen und somit Prozesse zu systematisieren, die ein Mensch intuitiv durchführt, etwa die Verknüpfung einer Aussage mit Erfahrungen aus der Vergangenheit, um sie semantisch korrekt einordnen zu können. Die Forschung in diesen Bereichen schreitet stetig fort, sodass Sprachassistenten künftig immer besser die Bedeutung aus einer menschlichen Aussage extrahieren können.

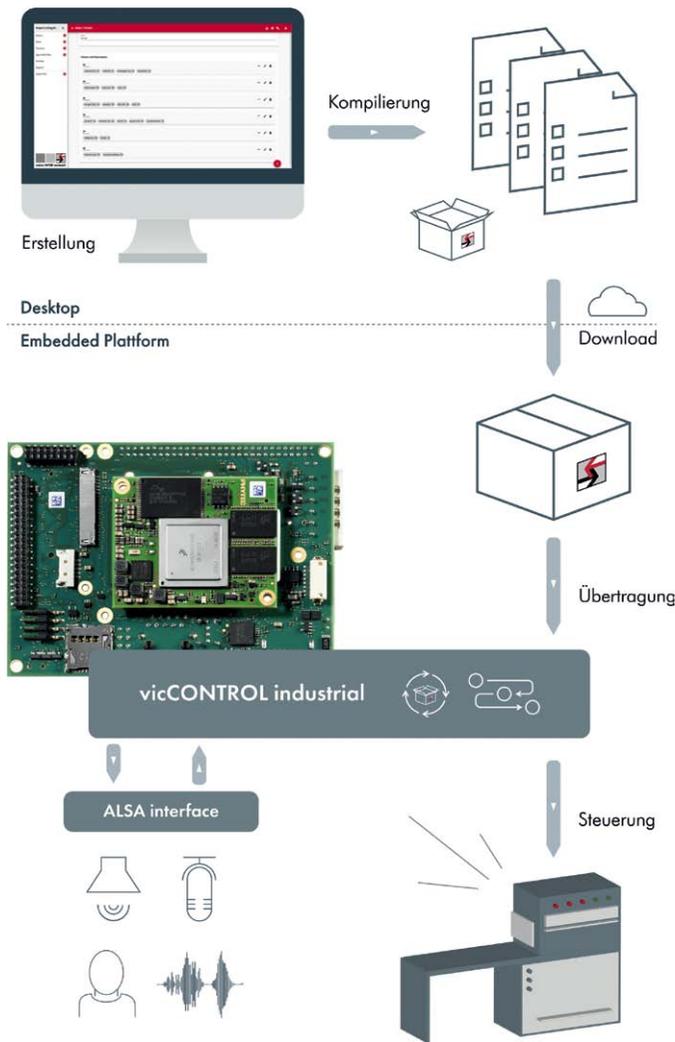
Wie groß ist derzeit die Nachfrage nach Sprachsteuerung in solchen Anwendungen? Wie wird sie sich Ihres Erachtens in absehbarer Zeit entwickeln?

Wenzel: Gerade durch Corona und das dadurch gesteigerte Bewusstsein für Hygiene stellen wir eine erhöhte Nachfrage nach berührungslosen Bedieneinheiten fest – ob im öffentlichen Raum, in der Gebäudeautomatisierung oder in Pflegeeinrichtungen. Dabei hat Sprachsteuerung aber noch weitere Vorteile: Sie ist barrierefrei, und die Bediener haben Hände und Augen frei für andere Aufgaben. Außerdem ist Sprachsteuerung sehr effizient, weil mit einem einzigen Befehl komplexe

Vorgänge ausgelöst werden können. Je einfacher die Lösung in industrielle Applikationen zu implementieren ist, umso schneller wird sich die Technologie verbreiten. Und genau daran arbeiten wir mit Lösungen wie vicControl industrial, das in über 30 Landessprachen erhältlich ist, schnell und problemlos in Anwendungen auf ARM- und x86-Basis integriert werden kann und mit Onboard-Datenverarbeitung und modernen Spracherkennungstechniken hohe Erkennungsraten bietet.

Welche Überlegungen sollten Unternehmen anstellen, bevor sie eine Sprachsteuerung in ihre Lösungen implementieren?

Wenzel: Ob der Einsatz von Sprachsteuerung sinnvoll oder sogar erforderlich sein könnte, lässt sich mit Überlegungen zur Nutzung prüfen: Lässt sich bei der Ein- und Ausgabe von Informationen Zeit sparen? Gibt es Vorgänge, die parallel ausgeführt werden könnten, für die aber keine Hände frei sind, oder gibt es



Mit vicControl industrial gelangen Kunden in wenigen Schritten zu einer integrierten Sprachsteuerung.

parallele Vorgänge, bei denen es dringend erforderlich wäre, Hände und/oder Augen zur Verfügung zu haben, etwa um Sicherheit zu gewährleisten? So lässt sich schnell herausfinden, ob Prozesse durch den Einsatz von Sprachassistenten vereinfacht und beschleunigt werden können. Sehr häufig stellen Unterneh-

men dabei fest, wie groß das Potenzial zur Effizienz- und Produktivitätssteigerung durch Sprachsteuerung tatsächlich ist. Gleichzeitig sind auch Gegenargumente zu prüfen – etwa ob die Umgebungslautstärke den Einsatz einer Sprachbedienung überhaupt zulässt oder ob steuerungsrelevante Daten über eine Abfrage erhoben werden können oder nicht. Daraus ergibt sich, welche Vor- und Nachteile Sprachsteuerung beispielsweise gegenüber einem Touchscreen hat – oder wie sich beide Bedientechniken optimal ergänzen.

Wie lassen sich Datenschutz und Eignung für die Industrie prüfen und umsetzen?

Wenzel: Tatsächlich ist Datenschutz das Hauptargument gegen den Einsatz der Technologie in der Industrie. Nur durch eine ausschließlich lokale Verarbeitung der Daten, also ohne Cloud-Anbindung,

lässt sich die Sicherheit von Kundeninformationen, eigenen Rezepturen oder Daten aus Entwicklung und Produktion gewährleisten. vicControl industrial arbeitet deshalb rein lokal. Eine Verknüpfung zur Cloud ist nicht nötig. Unternehmensdaten – angefangen mit Anzahl und Häufigkeit von Interaktionen – bleiben so

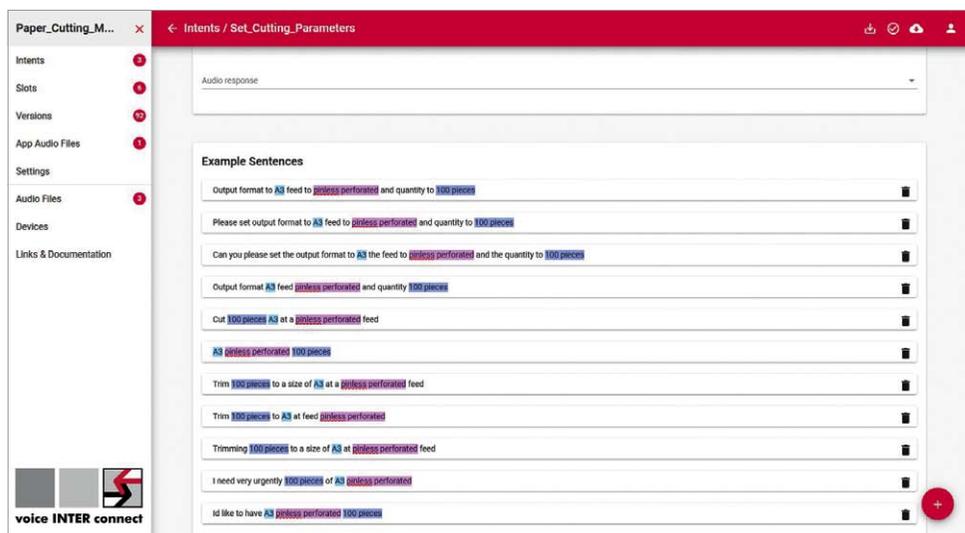
Was benötigen Unternehmen, um mit vicControl eigene Anwendungen zu erstellen? Wie funktioniert das?

Wenzel: vicControl industrial ist als Entwicklungskit mit diversen Peripheriebausteinen für Display, Audio und weitere Geräte erhältlich – alternativ auf Basis von ARM-basierten Embedded-Systemen oder x86-Rechnern, beides weit verbreitete industrielle Bedienplattformen. Außerdem ist durch Verknüpfung der Entwicklungskits mit vicSDC, einem umfangreichen Entwicklungs-Tool, die Erstellung und der Download kundenspezifischer Sprachsteuerungen stark vereinfacht. vicSDC ist mit seinem grafischen User Interface auch ohne tiefgehende Programmierkenntnisse bedienbar und bietet semantische Kategorien und Templates, etwa für Zeit- und Mengenangaben, sowie automatisierte Validierungsverfahren auf KI-Basis. Das reduziert das Risiko von Designfehlern. Für den Test sind Beispielapplikationen vorbereitet. Eine Vernetzung des Entwicklungskits und der Zielapplikation erfolgt über das IoT-Protokoll MQTT, sodass Nutzer schon mit dem Prototypen den Eindruck einer fertigen Bedienlösung erhalten. Durch kurze Iterationszyklen und Nutzertests lässt sich der Prototyp bis zur ergonomischen Perfektion verfeinern.

Die prototypisch auf Basis des Entwicklungskits erstellte Software-Lösung kann anschließend in das Design zahlreicher Geräte und Anwendungen übernommen werden.

Vor Kurzem hat voice Inter connect ein Modul für die Entwicklung von Kommunikationsgeräten auf SIP-Basis vorgestellt. Was hat es damit auf sich und wie passt das mit der Sprachsteuerung zusammen?

Diane Hirschfeld: Mit vicCom IP können Kunden rasch und problemlos individuelle Intercom-Lösungen entwickeln – ob für Notrufsäulen, Informations- und Kiosksysteme, für Pflegekraft-Rufsysteme oder für die Kollaboration in Industrie, Medizintechnik und vielen



Das umfangreiche Entwicklungs-Tool vicSDC mit grafischer Bedienoberfläche vereinfacht die Erstellung und den Download kundenspezifischer Sprachsteuerungen.



Martin Wenzel, voice Inter connect

„ Sprachassistenten werden künftig immer besser die Bedeutung aus einer menschlichen Aussage extrahieren können. “



Dr. Diane Hirschfeld, voice Inter connect

„ In Medizin und Pflege kann der Einsatz der Sprachbedienung in Rufsystemen Sicherheit, Komfort und Ergonomie für die Patienten verbessern. “

anderen Bereichen. Die Sprachbedienung von voice Inter connect ist ebenfalls bereits in das Modul integriert, sodass die Kommunikationsgeräte auf Wunsch die Vorteile der berührungsfreien Steuerung nutzen können – wie im eingangs genannten Fahrstuhl-Szenario.

Wie ist vicCom IP aufgebaut und wie wird es in Anwendungen integriert?

Hirschfeld: Primär für IP-Telefonie und Multimedia-Funktionen wie Durchsagen und Musik-Streaming ausgelegt, ist vicCom IP eine miniaturisierte Embedded-Plattform auf Basis eines ARM Cortex-A7. Mit PoE-Versorgung und zahlreichen Schnittstellen für Taster, LEDs, Keypad, Mikrofone und Lautsprecher ausgestattet, überlässt die Plattform Anwendern die Entscheidung, welches Gerät sie damit aufbauen wollen. So wird vicCom IP durch Anschluss einer LED und zweier 10-W-Lautsprecher zu einem IP-Verstärker. Durch Anschluss eines Keypads, einer LED, Lautsprecher und Handset sowie Verbau in einem robusten Metallgehäuse entsteht ein Industrie-Telefon. Anwender können die Benutzung der genannten Komponenten im Telefonie-Ablauf, den Einsatz von Licht- oder Tonsignalen oder die Priorisierung von Bedienelementen, etwa für Info- und Notrufe, über die integrierte Konfigurationsoberfläche (GUI) intuitiv und ohne Programmierkenntnisse festlegen.

Welche Funktion übernimmt dabei die Sprachbedienung?

Hirschfeld: Die Sprachbedienung ist zunächst mit dem Telefonbuch verknüpft, kann aber auch für beliebige Bedienaktionen als Eingabe konfiguriert werden und so die Wahl eines Rufziels oder einer Automatisierungsfunktion triggern. So können etwa mittels Ein-Knopf-Bedienung ein Rufziel aus einem 100 Einträge langen Telefonbuch ausgewählt oder Türen geöffnet und Beleuchtungen geschaltet werden.

Diese Funktionen sind mit herkömmlichen Bedienelementen aktuell nicht realisierbar.

Was verstehen Sie unter „(Industrie-) Intercom-Lösungen“, und welche Innovationen sind hier künftig zu erwarten?

Hirschfeld: Mit (Industrie-) Intercom-Lösungen haben wir unser eigenes Anwendungsfeld für Sprachbedienung – für den Einsatz überall dort, wo nicht nur das Bedürfnis nach Komfort befriedigt werden soll, sondern die Sprachbedienung professionellen Anwendern handfesten Mehrwert bietet. Wir haben noch viele weitere Ideen für clevere Automatisierungsfunktionen, etwa durch vereinfachte Konfiguration und Systemüberwachung oder allgegenwärtige Assistenzfunktionen.

Durch die Vernetzung mehrerer vicCom-IP-Plattformen über Ethernet entsteht ein lokales Netzwerk aus intelligenten Audiosensoren und Aktoren. Alle benötigten Elemente sind dabei in vicCom IP bereits vorhanden – von der Audioaufnahme in hoher Qualität über die IP-basierte Übertragung und Interpretation durch die Sprachbedienung bis zur Auslösung von Automatisierungsfunktionen mittels der integrierten I/Os. Dabei lassen sich die Komponenten drahtlos oder drahtgebunden vernetzen. Durch die Nutzung des SIP-Standards ist die vicCom-IP-Plattform mit klassischen IP- oder DECT-Telefonen kompatibel, also problemlos in vorhandene IP-Telefonie-Systeme migrierbar. Besonders in Medizin und Pflege kann der Einsatz der Sprachbedienung in Rufsystemen Sicherheit, Komfort und Ergonomie für die Patienten verbessern. Gleichzeitig trägt Spracherkennung dazu bei, die Effizienz der Pflegekräfte zu steigern und die Betreuung der Patienten zu verbessern, und zwar durch Parallelisierung von Arbeitsabläufen und Dokumentationspflichten.

Das Interview führte Andreas Knoll.