

**ZUR SOFORTIGEN VERÖFFENTLICHUNG**

**Nr. 3112**

*Bei diesem Text handelt es sich um eine Übersetzung der offiziellen englischen Version dieser Pressemitteilung, die nur als Hilfestellung und Referenz bereitgestellt wird. Ausführliche und/oder spezifische Informationen entnehmen Sie bitte der englischen Originalversion. Im Falle von Abweichungen hat der Inhalt der englischen Originalversion Vorrang.*

*Kundenanfragen*

Mitsubishi Electric Research Laboratories  
[www.MitsubishiElectric.com/ssl/contact/company/rd/form](http://www.MitsubishiElectric.com/ssl/contact/company/rd/form)  
[www.merl.com](http://www.merl.com)

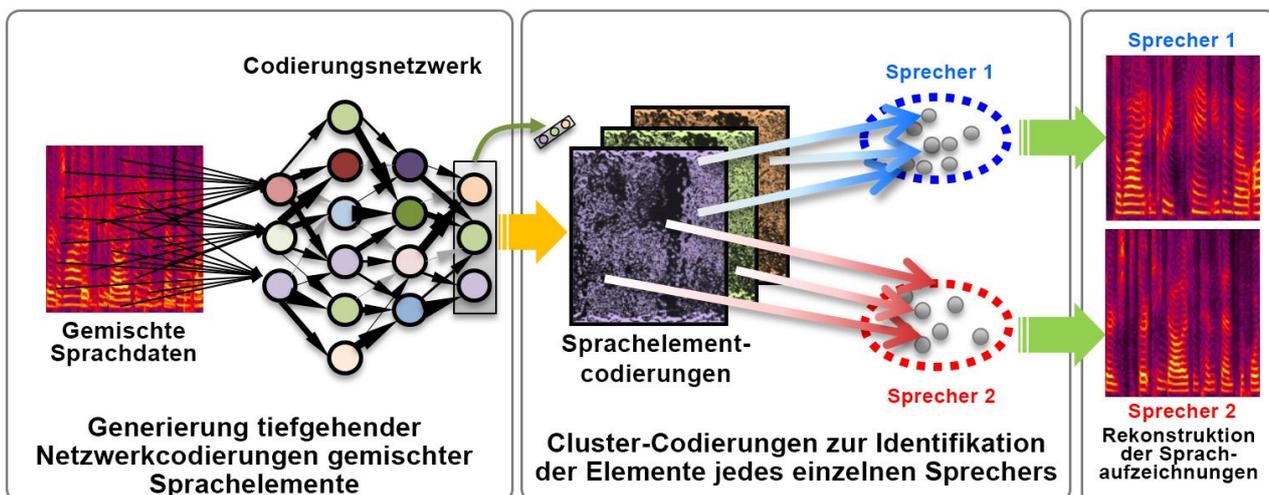
*Presseanfragen*

Public Relations Division  
Mitsubishi Electric Corporation  
[prd.gnews@nk.MitsubishiElectric.co.jp](mailto:prd.gnews@nk.MitsubishiElectric.co.jp)  
[www.MitsubishiElectric.com/news/](http://www.MitsubishiElectric.com/news/)

## **Mitsubishi Electric trennt gleichzeitige, mit einem Mikrofon aufgenommene Sprachaufzeichnungen mehrerer unbekannter Sprecher**

*Die Sprachtrennungstechnologie wird mit der proprietären, auf künstlicher Intelligenz basierenden „Deep Clustering“-Methode ermöglicht*

**TOKIO, 24. Mai 2017** – [Mitsubishi Electric Corporation](http://www.MitsubishiElectric.com) (TOKIO: 6503) gab heute die Entwicklung der weltweit ersten Technologie bekannt, die gleichzeitige, mit einem Mikrofon in Echtzeit aufgenommene Sprachaufzeichnungen mehrerer unbekannter Sprecher trennt und mit hoher Qualität rekonstruiert. Im Rahmen von Tests wurden die gleichzeitigen Sprachaufzeichnungen von zwei und drei Personen mit jeweils bis zu 90- bzw. 80-prozentiger Genauigkeit getrennt, was zum Zeitpunkt dieser Mitteilung nach Auffassung des Unternehmens eine absolute Weltneuheit ist. Die brandneue Technologie, die mit der proprietären, auf künstlicher Intelligenz (KI) basierenden „Deep Clustering“-Methode von Mitsubishi Electric ermöglicht wurde, soll zu einer besser verständlichen Sprachkommunikation und einer genaueren automatischen Spracherkennung beitragen.



Bei zwei gleichzeitigen Sprechern wurde eine Genauigkeit von über 90 Prozent erzielt, was im gewerblichen Bereich ausreicht. Mit herkömmlicher Technologie wurde im Vergleich dazu nur eine Genauigkeit von 51 Prozent erreicht. Die neue Technologie kann zwischen verschiedenen Kombinationen von mehreren gesprochenen Sprachen und Geschlecht unterscheiden. Die oben genannten Ergebnisse basieren auf optimalen Aufnahmebedingungen wie niedrigem Hintergrund-Geräuschpegel und Sprechern, die ungefähr gleich laut sprechen.

Bei der Deep Clustering-Technologie wird die proprietäre Deep Learning-Methode von Mitsubishi Electric eingesetzt, um zu ermitteln, wie Signalkomponenten der Originalsprachdaten mehrerer Sprecher codiert werden müssen, damit die Signalkomponenten jedes einzelnen Sprechers anhand ihrer Codierungen einfach voneinander unterschieden werden können. Um dies zu erreichen werden die Codierungen so optimiert, dass verschiedene Signalkomponenten desselben Sprechers ähnliche Codierungen aufweisen, während die Signalkomponenten anderer Sprecher völlig andere Codierungen haben. Die ermittelte Codierungstransformation wird auf die Eingabesprache angewendet, und die Codierungen der Signalkomponenten jedes Sprechers werden mit einem Clustering-Algorithmus identifiziert, der Datenpunkte in Abhängigkeit ihrer Gemeinsamkeiten in Gruppen verarbeitet. Anschließend wird die Sprachaufzeichnung jedes Sprechers durch Resynthese von dessen getrennten Sprachkomponenten rekonstruiert.

**Genauigkeit bei der Trennung gleichzeitiger Sprachaufzeichnungen mehrerer Sprecher\***

	Zwei Sprecher (ein Mikrofon)	Drei Sprecher (ein Mikrofon)
Neue Technologie	> 90 % (absolute Weltneuheit)	> 80 % (absolute Weltneuheit)
Herkömmliche Technologie	51 %	–

\* Auf der Basis optimaler Aufnahmebedingungen

Im gewerblichen Bereich (z. B. Freisprech-Telefonfunktionen von Fahrzeugnavigationssystemen) wird Geräuschunterdrückungstechnologie zur Beseitigung von Geräuschen aus mit Hintergrundgeräuschen gemischten Sprachaufzeichnungen eingesetzt. Um die Sprachaufzeichnung einer bestimmten Person von den gleichzeitigen Sprachaufzeichnungen anderer Personen zu trennen, sind bei den vorhandenen Methoden mehrere Mikrofone am Standort jedes Sprechers erforderlich, um die Informationen aufzuzeichnen. Bis jetzt gab es jedoch keine effektive Methode, um die mit nur einem Mikrofon aufgenommenen Sprachaufzeichnungen mehrerer unbekannter Sprecher genau zu rekonstruieren.

Als nächsten Schritt wird Mitsubishi Electric Möglichkeiten untersuchen, wie seine neue Technologie zur Verbesserung der Qualität der Sprachkommunikation und der Genauigkeit automatischer Spracherkennung in realen Umgebungen wie Fahrzeugen, Wohnungen und Aufzügen eingesetzt werden kann.

**Patente**

Angemeldete Patente für die in dieser Pressemitteilung bekannt gegebene Technologie: 8 außerhalb Japans.

###

### **Über die Mitsubishi Electric Corporation**

Mit über 90 Jahren Erfahrung in der Bereitstellung zuverlässiger, hochwertiger Produkte ist die Mitsubishi Electric Corporation (TOKIO: 6503) ein anerkanntes, weltweit führendes Unternehmen in der Herstellung, in der Vermarktung und im Vertrieb von Elektro- und Elektronikgeräten für die Informationsverarbeitung, Kommunikation, Raumfahrtentwicklung und Satellitenkommunikation, Unterhaltungselektronik, Industrietechnik, den Energie- und Transportsektor sowie Gebäudeanlagen. Im Sinne seiner Unternehmensphilosophie „Changes for the Better“ und Umwelterklärung „Eco Changes“ setzt sich Mitsubishi Electric als globales, im Umweltschutz führendes Unternehmen dafür ein, die Gesellschaft mit neuen Technologien zu bereichern. Das Unternehmen verzeichnete konzernweit einen konsolidierten Umsatz von 4.238,6 Mrd. Yen (37,8 Mrd. US-Dollar\*) im Geschäftsjahr zum 31. März 2017. Weitere Informationen erhalten Sie unter:

[www.MitsubishiElectric.com](http://www.MitsubishiElectric.com)

\* Zum Wechselkurs von 112 Yen für einen US-Dollar, der am 31. März 2017 von der Tokioter Devisenbörse angegeben wurde.