

MITSUBISHI ELECTRIC CORPORATION
PUBLIC RELATIONS DIVISION
7-3, Marunouchi 2-chome, Chiyoda-ku, Tokio 100-8310 (Japón)

PARA SU PUBLICACIÓN INMEDIATA

N.º 3112

Este texto es una traducción de la versión oficial en inglés de este comunicado de prensa y se le proporciona a modo de referencia, para su comodidad. Consulte el texto original en inglés para obtener detalles específicos. En caso de que ambas versiones difieran, prevalecerá el contenido de la versión en inglés.

Consultas de los clientes

Consultas de los medios

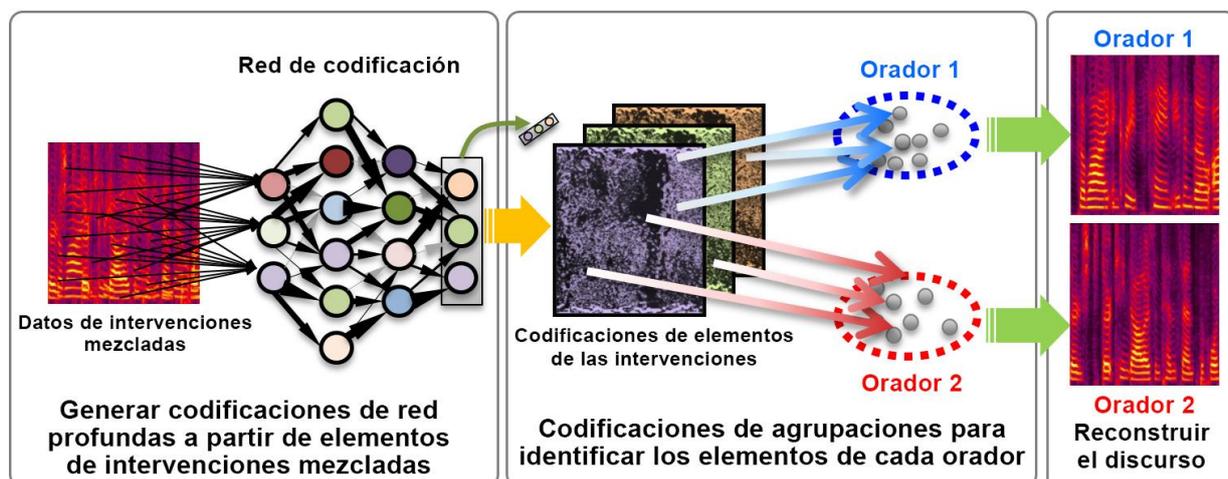
Mitsubishi Electric Research Laboratories
www.MitsubishiElectric.com/ssl/contact/company/rd/form
www.merl.com

Public Relations Division
Mitsubishi Electric Corporation
prd.gnews@nk.MitsubishiElectric.co.jp
www.MitsubishiElectric.com/news/

Mitsubishi Electric separa intervenciones simultáneas de varios oradores desconocidos grabados con un mismo micrófono

La tecnología de separación de discursos se ha logrado mediante el método de inteligencia artificial "Deep Clustering"

TOKIO, 24 de mayo de 2017 – [Mitsubishi Electric Corporation](http://www.MitsubishiElectric.com) (TOKIO: 6503) ha anunciado hoy que ha creado la primera tecnología del mundo que separa, y después reconstruye con calidad alta, la intervención simultánea de varios oradores desconocidos grabados con un único micrófono en tiempo real. En las pruebas, las intervenciones simultáneas de dos y de tres personas fueron separadas con una precisión de hasta un 90 % y un 80 %, respectivamente, lo que la empresa considera un anuncio sin precedentes de este tipo en el mundo. Se espera que esta novedosa tecnología, obtenida mediante el método de "Deep Clustering" (agrupación profunda) propiedad de Mitsubishi Electric y basado en inteligencia artificial (AI), contribuya a lograr comunicaciones de voz más inteligibles y un reconocimiento de voz automático más preciso.



En el caso de dos oradores simultáneos, la precisión superó el 90 %, lo que es suficiente para las aplicaciones comerciales, en comparación con el 51 % de precisión mediante la tecnología convencional. La nueva tecnología es capaz de distinguir entre combinaciones de varios idiomas hablados y de género. Los resultados anteriores se basan en condiciones de grabación ideales, incluidos niveles bajos de ruido ambiente y oradores que se encuentran hablando a un volumen muy similar.

La tecnología Deep Clustering utiliza el método de aprendizaje profundo, propiedad de Mitsubishi Electric, para aprender a codificar los datos de los componentes de la señal de la intervención original de varias personas, para que se puedan distinguir fácilmente los componentes de la señal que pertenecen a cada orador por sus codificaciones. Para lograr esto, se optimizan las codificaciones para que los diferentes componentes de la señal que pertenecen al mismo orador tengan codificaciones similares y los que pertenecen a los diferentes oradores tengan codificaciones diferentes. La transformación de codificación aprendida se aplica al discurso de entrada y se identifican las codificaciones de los componentes de la señal de cada orador mediante un algoritmo de agrupación que procesa los puntos de datos en grupos, según sus similitudes. A continuación, se reconstruye la intervención de cada persona mediante la resintetización de los componentes de la intervención por separado.

Precisión en la separación de intervenciones simultáneas de múltiples oradores*

	Dos oradores (un solo micrófono)	Tres oradores (un solo micrófono)
Nueva tecnología	>90 % (sin precedentes)	>80 % (sin precedentes)
Tecnología tradicional	51 %	—

*Basado en condiciones de grabación ideales

Esta tecnología de reducción del ruido, utilizada para eliminar el ruido de una mezcla de discurso y ruido, se implementa en aplicaciones comerciales como las funciones de telefonía de manos libres de los sistemas de navegación para automóviles. Para separar la intervención de una persona objetivo de la intervención simultánea de otras personas, los métodos existentes requieren que varios micrófonos obtengan información sobre la ubicación de cada orador. Sin embargo, hasta ahora no ha habido ningún método eficaz para reconstruir con precisión la intervención de varios oradores desconocidos grabados mediante un único micrófono.

En adelante, Mitsubishi Electric explorará las posibles oportunidades para aplicar su nueva tecnología y mejorar la calidad de las comunicaciones de voz y la precisión del reconocimiento automático de las intervenciones en entornos reales, como coches, viviendas o ascensores.

Patentes

Patentes pendientes de la tecnología anunciada en este comunicado de prensa: ocho en el extranjero.

###

Acerca de Mitsubishi Electric Corporation

Con más de 90 años de experiencia en el suministro de productos fiables y de alta calidad, Mitsubishi Electric Corporation (TOKIO: 6503) es un líder mundial reconocido en la fabricación, comercialización y venta de equipos eléctricos y electrónicos utilizados en el procesamiento de la información y las comunicaciones, en el desarrollo espacial y las comunicaciones por satélite, en los aparatos electrónicos de consumo, en la tecnología industrial, en la energía, en el transporte y en los equipos de construcción. Aprovechando el espíritu de su declaración corporativa "Changes for the Better" y su declaración medioambiental "Eco Changes", Mitsubishi Electric se esfuerza por ser una empresa internacional comprometida con el medio ambiente líder y por enriquecer la sociedad con la tecnología. La empresa registró ventas de grupo consolidadas de 4 238,6 mil millones de yenes (unos 37,8 mil millones de dólares estadounidenses*) en el ejercicio fiscal que terminó el 31 de marzo de 2017. Para obtener más información, visite:

www.MitsubishiElectric.com

*Tipo de cambio de 112 yenes por dólar estadounidense, tipo concedido por el Mercado de divisas de Tokio el 31 de marzo de 2017