



INTRODUCCIÓN

Los sistemas de reconocimiento del habla en la actualidad, aunque son capaces de ofrecer buenos resultados en condiciones favorables, disminuyen sus prestaciones enormemente cuando se enfrentan a situaciones de aplicación real. Esto es fundamentalmente debido a la influencia del ruido y de la reverberación.

Estos dos factores degradantes son especialmente significativos cuando empleamos **micrófonos distantes**. Esta situación es muy habitual en diferentes aplicaciones como salas de reuniones, videoconferencias o incluso en el interior de vehículos, donde el locutor no puede disponer de un micrófono de solapa.

Una de las posibilidades de emergente interés para resolver el problema del reconocimiento con micrófonos lejanos es el uso de técnicas **multi-micrófono** que permitan mejorar la calidad de la señal recibida mediante la adecuada combinación o uso de varios micrófonos. Además, el uso de varios micrófonos no sólo es útil para mejorar las prestaciones de los sistemas de reconocimiento, sino que en general puede ser útil para la mejora de cualquier señal sonora, como por ejemplo, el uso en aplicaciones biomédicas como los implantes cocleares.



Aunque en principio estas técnicas ofrecen prometedoras expectativas, también presentan numerosas dificultades. Uno de estos inconvenientes que presentan muchas de las técnicas multi-micrófono, es la necesidad de conocer previamente la posición de la fuente de voz. Por ello, es fundamental para el éxito de las técnicas multi-micrófono disponer de técnicas robustas que nos permitan estimar adecuadamente la posición del hablante, en definitiva, disponer de algoritmos que nos permitan la **localización de fuentes sonoras** de una forma fiable.

OBJETIVOS

- Trabajo bibliográfico. Conocer y entender los métodos actuales de localización de fuentes mediante técnicas multi-micrófono y los problemas que entraña.
- Trabajo de implementación. Desarrollar los métodos de localización de fuentes sonoras que se hayan considerado más relevantes basándose en el estudio del estado del arte actual.
- Trabajo experimental y de comparación. Comparar mediante experimentos simulados y reales los métodos seleccionados. Hacer un pequeño estudio del impacto de la estimación errónea de DOA en un conformador sencillo de micrófonos destinado al reconocimiento del habla.
- Trabajo adicional. Con la experiencia obtenida proponer alguna modificación o alguna aportación a las técnicas bajo estudio.
- ...

PERFIL DE LOS CANDIDATOS

- Interés por el procesamiento de señal.
- Haber cursado la asignatura optativa Procesado de Voz.
- Haber cursado la asignatura optativa Arrays.
- Buen manejo de los lenguajes de programación Matlab y C/C++.

CONTACTO

Alberto Abad
TALP Research Center
Department of Signal Theory and Communications
D5-118, Campus Nord, UPC
alberto@gps.tsc.upc.es