

Plataforma acústica de bajo coste basada en arrays de micrófonos MEMS para la detección y localización de blancos

Izquierdo Fuente, Alberto ^{1,*}, Villacorta Calvo, Juan José ¹, del Val Puente, Lara ¹ y Alloza, Pablo ¹.

¹ Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática. ETSIT. Universidad de Valladolid. Paseo Belén 15, 47011, Valladolid (España). Correo electrónico: (AIF) alberto.izquierdo@uva.es, (JVC) juavil@tel.uva.es, (LVP) lara.val@uva.es, (PA) pablo.alloza@alumnos.uva.es

* Autor Principal y responsable del trabajo; Correo electrónico: (AIF) alberto.izquierdo@uva.es

Resumen: Este trabajo presenta una arquitectura de adquisición y procesado de señales acústicas que, utilizando un array de micrófonos MEMS (*Micro-Electro-Mechanical Systems*) y sistemas mixtos de procesado con procesadores multicore y FPGAs (*Field Programmable Gate Array*), obtiene imágenes acústicas de alta resolución. El sistema permite, mediante técnicas de beamforming, conocer la posición espacial de fuentes acústicas (trabajando en modo pasivo, es decir, en modo escucha) y la posición de los objetos en un espacio de vigilancia (trabajando en modo activo, es decir, en modo radar acústico). Las aplicaciones de este sistema son múltiples, desde la obtención de mapas de ruido, hasta la localización de objetos. Adicionalmente, el trabajo incluye un conjunto de casos de estudio para las principales aplicaciones de uso para ambos modos de funcionamiento.

Palabras clave: plataforma acústica, bajo coste, array de micrófonos MEMS, detección, localización

1. Introducción

Hoy en día, las imágenes acústicas están mostrando un gran desarrollo. En la actualidad, las imágenes acústicas están asociadas a una gran variedad de aplicaciones [1]. Algunas de estas aplicaciones se basan en sistemas pasivos, que escuchan el entorno para localizar los sonidos. Otras aplicaciones se basan en sistemas activos, según el principio SODAR (*Sound Detection And Ranging*). Estos sistemas activos envían una señal acústica y esperan su reflejo en posibles objetos para detectarlos. Muchos de estos sistemas se basan en el uso de arrays.

Un array es un conjunto ordenado de sensores idénticos, excitados de una manera específica [2]. Mediante el uso de técnicas de conformación o *beamforming* [3], el patrón del haz del conjunto se dirige electrónicamente hacia diferentes direcciones, lo que permite la discriminación de las fuentes acústicas en función de su posición. En concreto, los arrays de micrófonos MEMS (*Micro-Electro-Mechanical Systems*) están compuestos por micrófonos de alta calidad con una elevada SNR (relación señal/ruido), bajo consumo de energía y alta sensibilidad [4]. Los autores de este trabajo tienen experiencia en el diseño [5] y desarrollo de sistemas de imagen acústica, basados en arrays, para su aplicación en sistemas de vigilancia [6], ambientes asistidos [7], sistemas de identificación biométrica [8], detección de fallos en maquinaria [9], detección de peatones [10], o detección desde un dron de señales de auxilio [11].

Este trabajo presenta un sistema que adquiere y procesa señales para obtener imágenes acústicas. La tarea de adquisición del sistema puede funcionar en dos modos: i) modo pasivo (es decir, modo de escucha), y ii) modo activo (es decir, modo SODAR). Éste es un sistema multiplataforma, en el que cada tarea de procesamiento puede intercambiarse entre sus distintos niveles. Así, el sistema puede adaptarse a diferentes escenarios de coste y movilidad gracias a su característica reconfigurable. Dado que las aplicaciones de este sistema son múltiples, este documento también incluye casos de uso para ambos modos de funcionamiento.

2. Descripción del sistema

2.1. Configuración del hardware

2.1.1. Sistema de procesado

Se ha seleccionado una plataforma sbRIO 96xx [12] como unidad base para este sistema. Esta plataforma pertenece a la familia de dispositivos RIO (*Reconfigurable Input-Output*) de National Instruments que está orientada a sensores con procedimientos de adquisición no estándar, permitiendo la programación a bajo nivel de las rutinas de adquisición. Se seleccionó una solución comercial para reducir los costes, ya que una solución basada en un hardware diseñado específicamente sería más cara. En concreto, la plataforma sbRIO es un controlador integrado de placa única que incorpora una FPGA (*Field Programmable Gate Array*) de Xilinx y un procesador multinúcleo. El procesador integrado incluido en una sbRIO es capaz de ejecutar todos los algoritmos de software para generar las imágenes acústicas, por lo que puede utilizarse como un módulo de array independiente formado por una sbRIO conectada a una placa de array MEMS tal y como se muestra en la Figura 1.

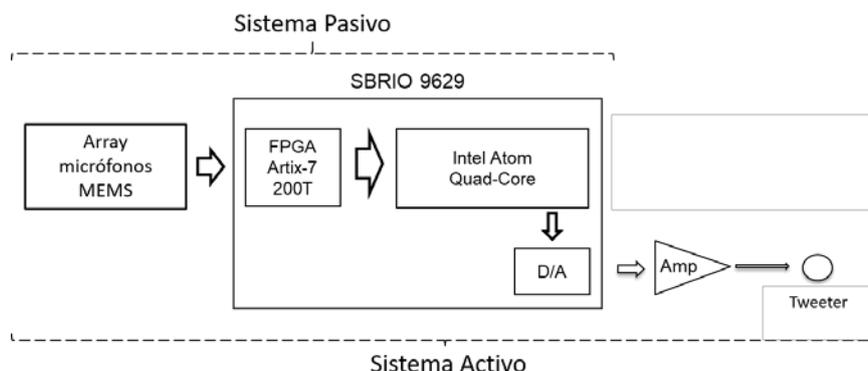


Figura 1. Diagrama de configuración del hardware.

2.1.2. Array MEMS

El sistema de adquisición de imágenes acústicas que se presenta en este trabajo se basa en arrays planares uniformes (UPA – *Uniform Planar Array*) de micrófonos MEMS [13], con el fin de obtener información de posición tridimensional. Los micrófonos digitales MEMS incluyen un micrófono, un acondicionador de señal y un dispositivo de adquisición incorporados en el propio chip. Por este motivo, un sistema de adquisición y procesamiento basado en arrays de micrófonos MEMS se reduce a dos elementos básicos: Micrófonos MEMS y un sistema de procesamiento. La integración del preamplificador del micrófono y el convertor AD (Analógico-Digital) en un solo chip reduce significativamente los costes de un sistema acústico de adquisición, si se compara con las soluciones basadas en micrófonos analógicos. Esta tecnología también reduce el espacio ocupado por el sistema, lo que hace factible la construcción de arrays con cientos o incluso miles de sensores.

2.2. Algoritmos software

Los algoritmos implementados en el sistema, mostrados en la Figura 2, pueden dividirse en cuatro bloques: Transmisión, Adquisición MEMS, Procesado de señales y Generación de imágenes. La Figura 2a muestra los bloques operativos en el modo de trabajo pasivo del sistema, y la Figura 2b en el modo activo.

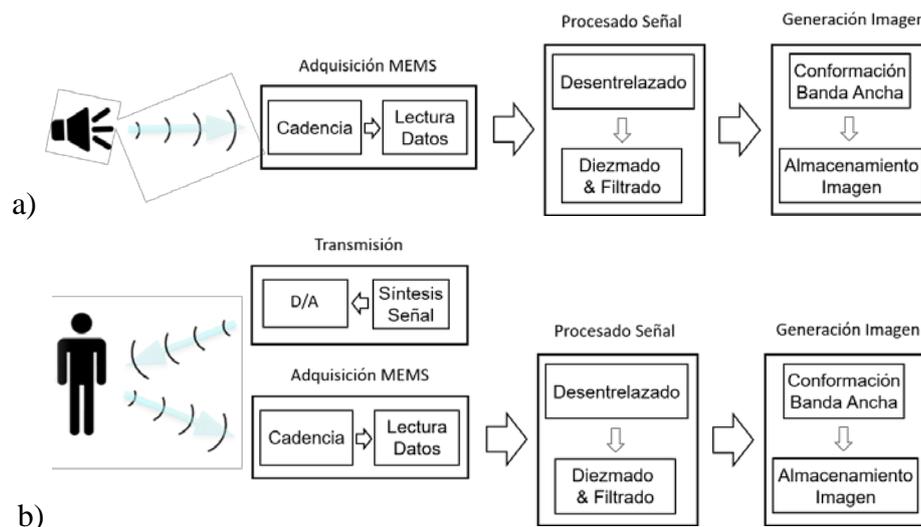


Figura 2. Diagrama de algoritmos software. a) Modo pasivo. b) Modo activo.

- 1) El bloque de Transmisión genera una señal multitono pulsada a través del convertidor DA (Digital-Analógico) hacia el amplificador de señal, y de ahí al altavoz, que transmite la señal. Como se puede observar en la Figura 2, este bloque sólo funciona si el sistema está trabajando en su modo activo.
- 2) En el bloque de Adquisición, cada micrófono MEMS adquiere la señal acústica. Como se observa en la Figura 2, si el sistema está trabajando en el modo pasivo, la señal acústica que adquieren los micrófonos MEMS es generada por el entorno en el que está trabajando el sistema. Por lo tanto, si el sistema trabaja en modo activo, la señal acústica que adquiere el micrófono es la reflexión de la señal transmitida por el propio sistema. Este bloque de Adquisición se implementa en la FPGA, generando una señal de reloj común para todos los MEMS, y leyendo simultáneamente todas las señales de los sensores a través de las entradas digitales de la FPGA.

- 3) En el bloque de Procesamiento de Señales se implementan dos rutinas:
 - a. Desentrelazado: Mediante este proceso se extrae la señal de un bit de cada micrófono MEMS de cada palabra binaria.
 - b. Diezmado y filtrado: Aplicando técnicas de *downsampling*, basadas en el diezmado y el filtrado, se obtiene una señal independiente de cada micrófono MEMS y se reduce la frecuencia de muestreo.
- 4) Por último, en el bloque de Generación de Imágenes, basado en la conformación de banda ancha, se define un conjunto de $N \times N$ direcciones de apuntamiento, y se evalúa la salida del conformador para cada una de estas direcciones. A continuación, las imágenes generadas se visualizan y se almacenan en el sistema.

3. Casos de estudio

3.1. Caso A: Modo pasivo

Este caso de estudio presenta la adquisición de imágenes acústicas de una matriz de ventiladores, con el fin de detectar si alguno de ellos no estaba funcionando correctamente.

En este trabajo, el array MEMS específico que se utilizó estaba compuesto por 4 módulos UPA de 8×8 micrófonos MEMS digitales separados entre sí una distancia uniforme de 2.5cm. Así, la adquisición de la señal acústica se realiza mediante un array global de 16×16 , es decir, de 256 micrófonos MEMS. Este sistema de adquisición obtuvo imágenes acústicas, en acimut y elevación, de una matriz 3×3 de ventiladores. Cada uno de los ventiladores empleados para construir la matriz era un Foxconn D90SM-12 3-Pin con 7 aspas. Los ventiladores de la matriz estaban controlados por una placa de interfaz de relés de 8 canales de Kkmoon, que permitía encender y apagar cada uno de los ventiladores de forma independiente. Como se muestra en la Figura 3, para las pruebas, la matriz de ventiladores se colocó a 50cm frente al array, dentro de una cámara anecoica de $5\text{m} \times 3\text{m} \times 2.5\text{m}$.

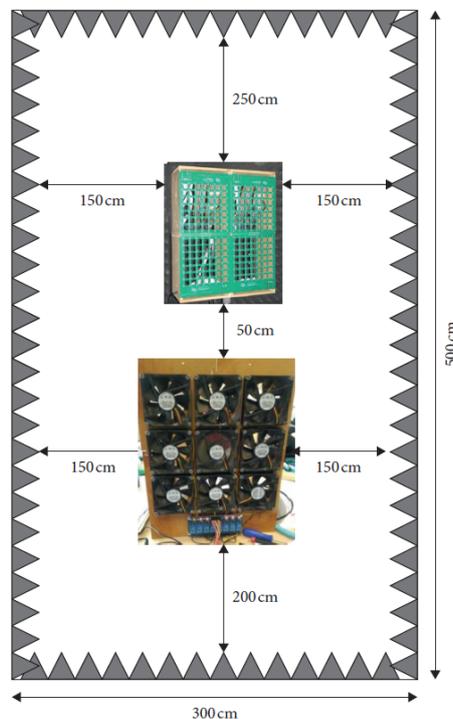


Figura 3. Diagrama de algoritmos software. a) Modo pasivo. b) Modo activo.

En este caso de estudio, las pruebas simulaban una matriz defectuosa en la que sólo funcionaba uno de los ventiladores. Las imágenes acústicas de la matriz funcionando en estas circunstancias específicas, es decir, con un único ventilador funcionando en cada caso, se muestran en la Figura 4. Las imágenes acústicas de cada uno de los 9 ventiladores de la matriz funcionando de forma única se obtuvieron con el uso del array de micrófonos MEMS junto con técnicas de *beamforming* de banda ancha, utilizando una frecuencia de trabajo de 1100Hz. Para cada ventilador se generaron 1500 imágenes acústicas, con el fin de obtener 9 imágenes acústicas promediadas, que se muestran en la Figura 4. Analizando estas imágenes, se puede observar que el centro de cada imagen parece estar situado en la misma posición de la matriz donde se encuentra el ventilador real, mostrado en la figura con una 'x'.

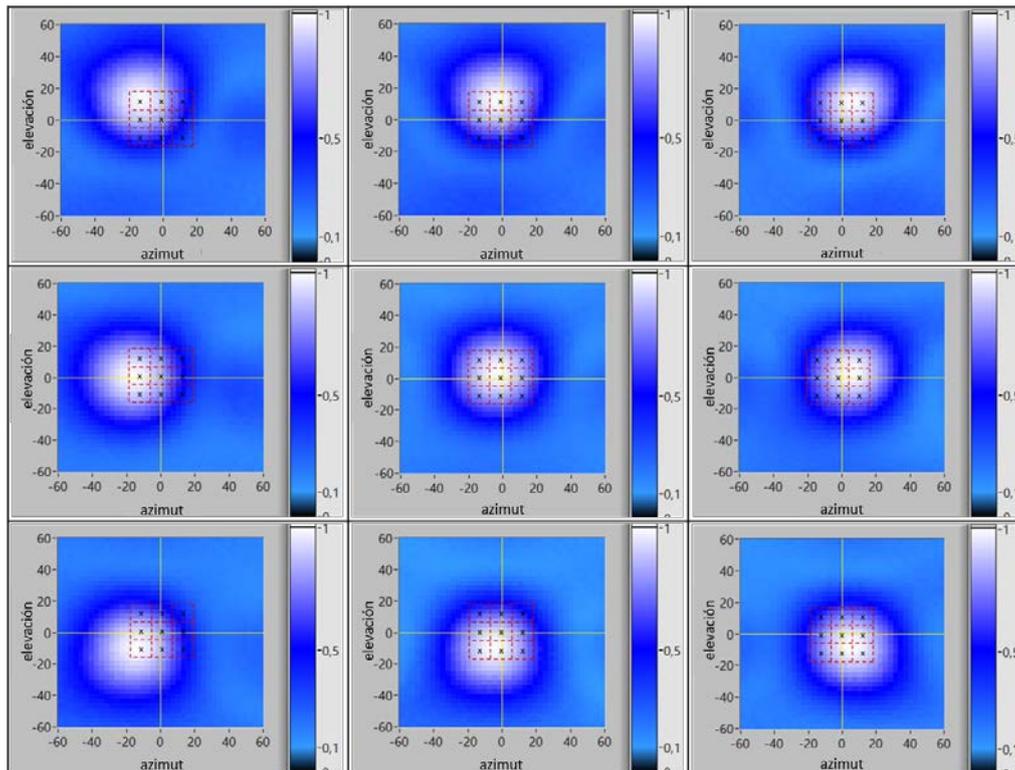


Figura 4. Imágenes acústicas de una matriz de ventiladores con un único ventilador en funcionamiento.

Para cada ventilador, el centro de cada una de las 1500 imágenes acústicas utilizadas para obtener la promediada se muestran en la Figura 5a, donde se puede observar que los centros de las imágenes acústicas de cada ventilador están dispersos alrededor del correspondiente "centro promediado". Gracias a estos "centros promediados", se han podido estimar las posiciones de los ventiladores en la matriz.

En la Figura 5b, los "centros promediados" pueden compararse con las posiciones reales de los ventiladores. Se puede observar que las posiciones estimadas de los centros son bastante cercanas a las reales, con errores inferiores a 2.5° . Por lo tanto, las imágenes acústicas promediadas de los ventiladores podrían utilizarse para estimar sus posiciones.

Los parámetros geométricos de estas imágenes acústicas para diferentes frecuencias se utilizaron después para entrenar un algoritmo de aprendizaje automático, en particular un clasificador de Máquinas de Vectores de Soporte (SVM – *Support Vector Machine*), con el fin de detectar el ventilador defectuoso.

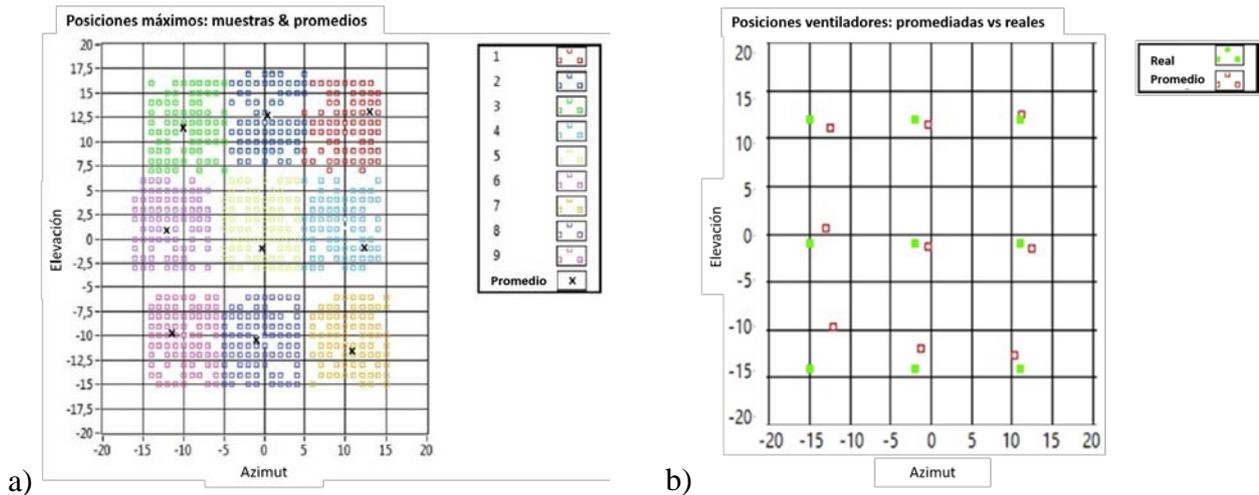


Figura 5. (a) Centros calculados para las imágenes acústicas de cada ventilador de la matriz funcionando de forma única. (b) Coordenadas del centro real vs. estimado.

3.2. Caso B: Modo activo

Este caso de estudio presenta la adquisición de imágenes acústicas radar, en el entorno delante de un vehículo, con el fin de detectar si hay peatones en su trayectoria.

En este trabajo, el array MEMS específico que se utilizó estaba compuesto por 6 módulos UPA de 5×5 micrófonos digitales MEMS uniformemente espaciados cada 0.9cm. Por lo tanto, la adquisición de la señal acústica se realiza mediante un array de 5×30 , es decir, mediante 150 micrófonos MEMS. En este caso, la dimensión horizontal del array era mayor que la vertical, ya que las coordenadas de interés de la posición del peatón eran el acimut y el rango. En este caso, la información de elevación, no era relevante.

Para las pruebas, la señal transmitida empleada fue un pulso de 3ms de 20kHz, generada por un tweeter. Se empleó además un vehículo compacto, y el escenario de las pruebas consistió en una carretera normal de 4m de ancho, con farolas y árboles a lo largo de sus bordes, como se muestra en la Tabla 1. En el escenario de pruebas, también se colocó una persona a 10m de distancia del sistema de captura, como puede observarse en la Figura 6.

ID	Objeto	Rango [m]	Acimut [°]
T1	Árbol	7.8	-51
P	Peatón	10.5	3
T2	Árbol	12.6	-28
L1	Farola	12.8	42
B	Papelera	14.4	19
T3	Árbol	18.3	-18
L2	Farola	21.3	-15
L3	Farola	22.7	26
T4	Árbol	24.3	-13
T5	Árbol	24.4	-25

Tabla 1. Posición de los objetos en el escenario de pruebas.

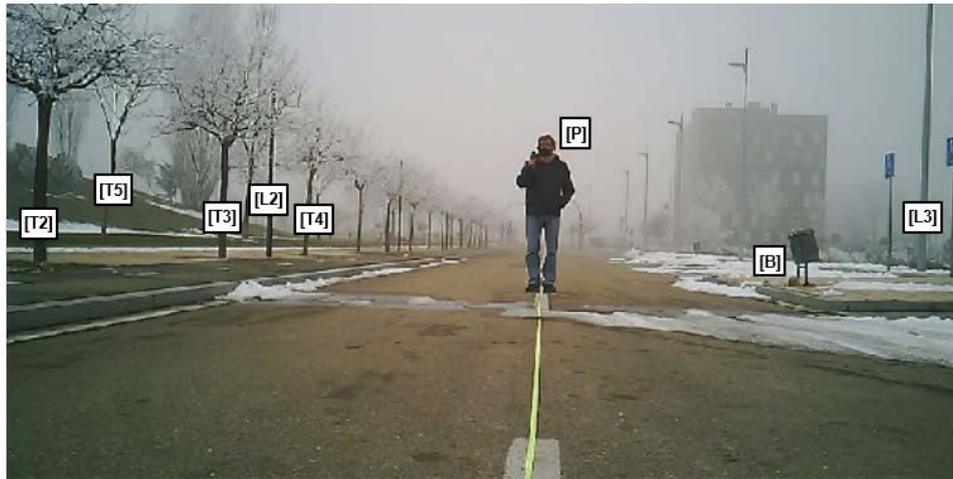


Figura 6. Escenario de pruebas.

La Figura 7 representa un ejemplo de las imágenes acústicas obtenidas en el espacio acimut/rango. En la imagen, las cruces rojas representan los objetivos detectados. Es evidente que el sistema detectó tanto a la persona, situada a unos 10m, como a otros objetos situados delante de ella, como las farolas y los árboles del borde de la carretera. En la figura también se pueden visualizar los límites del carril considerado, representados por una línea verde discontinua. La mayoría de los objetivos detectados estaban fuera del carril de la carretera y, por tanto, no debían tenerse en cuenta en el trabajo de detección posterior. En este caso, sólo el peatón quedaría como objetivo detectado.

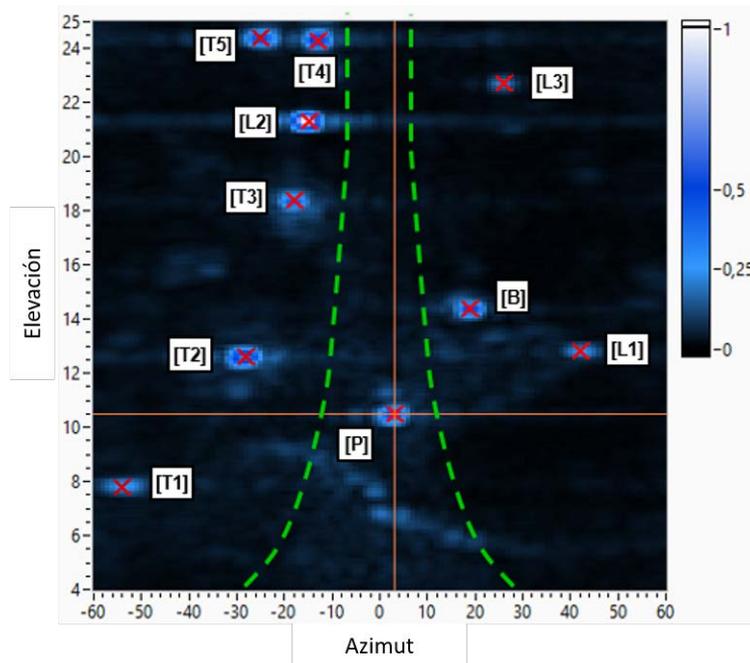


Figura 7. Imagen acústica con objetos detectados.

4. Conclusiones

Este trabajo presenta un sistema de adquisición y procesamiento de señales que obtiene imágenes acústicas de alta resolución mediante técnicas de conformación o *beamforming*. El sistema de adquisición acústica está basado en grandes arrays de micrófonos MEMS. Por otro lado, la parte de procesamiento del sistema se basa en procesadores multinúcleo y FPGAs.

Con algunos ejemplos se ha demostrado que el sistema puede funcionar tanto en modo pasivo como activo. El sistema es capaz de obtener la posición espacial de las fuentes acústicas, tales como gritos de auxilio o disparos, cuando trabaja en modo pasivo, y la posición de los objetos o de intrusos en un espacio de vigilancia si trabaja en modo activo.

La ventaja que aporta un radar acústico en este tipo de aplicaciones es que puede funcionar adecuadamente en entornos con visibilidad reducida, tales como entornos oscuros, o con niebla, lluvia copiosa o humo. Este sistema podría ser un complemento perfecto a cámaras de seguridad o a otros sistemas de detección, en las situaciones previamente mencionadas.

Referencias

1. Gan, W.S. *Acoustic Imaging: Techniques and Applications for Engineers*. John Wiley & Sons, **2012**.
2. Van Trees, H. *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*. John Wiley & Sons, **2002**.
3. Van Veen, B.D.; Buckley, K.M. Beamforming: A Versatile Approach to Spatial Filtering. **IEEE ASSP Magazine**, **1988**,5(2):4–24.
4. Scheeper, P.R.; van der Donk, A.G.H.; Olthuis, W.; Bergveld, P. A review of silicon microphones. *Sensors and Actuators A: Physical*. **1994**,44:1–11.
5. Del Val, L.; Jiménez, M.; Izquierdo, A.; Villacorta, J. Optimisation of sensor positions in random linear arrays based on statistical relations between geometry and performance. *Applied Acoustics*. **2012**,73(1):78–82.
6. Izquierdo, A.; Villacorta, J.; Raboso, M.; Martínez, A.; Rodríguez, D.; del Val, L. A human classification system for a video-acoustic detection platform. In the *International Carnahan Conference on Security Technology*, Albuquerque, NM, USA, 12–15 October **2004**. Available from: <https://ieeexplore.ieee.org/document/1405383>.
7. Villacorta-Calvo, J.; Jiménez-Gómez, M.I.; del Val-Puente, L.; Izquierdo-Fuente, A. A Configurable Sensor Network Applied to Ambient Assisted Living. *Sensors*. **2011**,11(11):10724–1073.
8. Del Val, L.; Izquierdo, A.; Villacorta, J.; Raboso, M. Acoustic Biometric System Based on Preprocessing Techniques and Support Vector Machines. *Sensors*. **2015**,15(6):14241–14260.
9. Del Val, L.; Izquierdo, A.; Villacorta, J.J.; Suárez, L. Comparison of Methodologies for the Detection of Multiple Failures Using Acoustic Images in Fan Matrices. *Shock and Vibration*. **2020**,2020:5816050.
10. Izquierdo, A.; Del Val, L.; Villacorta, J.J. Feasibility of Using a MEMS Microphone Array for Pedestrian Detection in an Autonomous Emergency Braking System. *Sensors*, **2021**,21(12):4162.
11. Izquierdo, A.; Del Val, L.; Villacorta, J.J.; Zhen, W.; Scherer, S.; Fang, Z. Feasibility of Discriminating UAV Propellers Noise from Distress Signals to Locate People in Enclosed Environments Using MEMS Microphone Arrays. *Sensors*. **2020**;20(3):597.
12. sbRIO-96xx Platform. [Internet]. [cited 2022 Jul 15]. Available from: <https://www.ni.com/docs/en-US/bundle/combo-sbrio-getting-started/page/um-overview.html>.
13. SPH0641LU4H-1 MEMS microphone. [Internet]. [cited 2022 Aug 16]. Available from: <https://www.knowles.com/docs/default-source/model-downloads/sph0641lu4h-1-revb.pdf>.