Design and Evaluation of a Scalable and Reconfigurable Multi-Platform System for Acoustic Imaging

Alberto Izquierdo [1,*], Juan José Villacorta [1], Lara del Val Puente [2] and Luis Suárez [3]

[1] Signal Theory and Communications Department, University of Valladolid, Valladolid 47011, Spain; juavil@tel.uva.es
[2] Mechanical Engineering Area, Industrial Engineering School, University of Valladolid, Valladolid 47011, Spain; lvalpue@eii.uva.es
[3] Civil Engineering Department, Superior Technical College, University of Burgos, Burgos 09001, Spain; luis.a.suarez.vivar@gmail.com
[*] Correspondence: alberto.izquierdo@tel.uva.es; Tel.: +34-983-185-801

Abstract: This paper proposes a scalable and multi-platform framework for signal acquisition and processing, which allows for the generation of acoustic images using planar arrays of MEMS (Micro-Electro-Mechanical Systems) microphones with low development and deployment costs. Acoustic characterization of MEMS sensors was performed, and the beam pattern of a module, based on an 8×8 planar array and of several clusters of modules, was obtained. A flexible framework, formed by an FPGA, an embedded processor, a computer desktop, and a graphic processing unit, was defined. The processing times of the algorithms used to obtain the acoustic images, including signal processing and wideband beamforming via FFT, were evaluated in each subsystem of the framework. Based on this analysis, three frameworks are proposed, defined by the specific subsystems used and the algorithms shared. Finally, a set of acoustic images obtained from sound reflected from a person are presented as a case study in the field of biometric identification. These results reveal the feasibility of the proposed system.

Keywords: MEMS array; scalable system; multi-platform framework; acoustic imaging

1. Introduction

In recent years, techniques for obtaining acoustic images have been developed rapidly. At present, acoustic images are associated with a wide variety of applications [1], such as non-destructive testing of materials, medical imaging, underwater imaging, SONAR, geophysical exploration, etc. These techniques for obtaining acoustic images are based on the RADAR (RAdio Detection and Ranging) principles, which form an image of an object from the radio waves that have been reflected on it [2]. RADAR systems require high-cost hardware and their application with people and specific materials is difficult, due to their low reflectivity. These are the reasons why acoustic imaging techniques, also called SODAR (Sound Detection and Ranging) techniques, were developed. These SODAR techniques, mainly based on the use of arrays, represent a simple, low-cost alternative for obtaining "acoustic images" of an object.

An array is an arranged set of identical sensors, excited in a specific manner [3]. Microphone arrays are a particular case, used in applications such as speech processing, echo cancellation, localization, and sound sources separation [4]. By using beamforming techniques [5], the array beam pattern can be electronically steered to different spatial positions, allowing the discrimination of acoustic sources based on their position.

The authors of this paper have experience in the design [6] and development of acoustic imaging systems, based in acoustic arrays, used in many different fields, such as detection and tracking systems [7,8], Ambient Assisted Living [9], or biometric identification systems [10–12]. The arrays used were ULA (Uniform Linear Array), formed by acoustic sensors distributed uniformly along a line. These arrays are simple, but limited to one dimension (azimuth or elevation) to estimate the spatial position of the sound source. In order to obtain spatial information in two dimensions, it is necessary to work with planar arrays with sensors distributed on a surface.

Working with planar arrays leads to an increase in both system complexity and space required by the acoustic sensors and the associated hardware, since the extension from a 1D array of N element, to a 2D array of N × N elements increases the number of the required channels in an order of N^2. These factors are accentuated when the number of sensors to work with is large. In the design of a system for acquiring and processing signals from an acoustic array, it should be noted that costs and complexity are directly related to the number of channels/sensors of the system. A typical system to obtain acoustic images has four basic elements: sensors, signal conditioners, acquisition devices and signal processor. For the first three elements, system cost increases linearly with the number of channels, as each sensor needs a signal conditioner and an acquisition device.

Digital MEMS (Micro-Electro-Mechanical System) microphones include a microphone, a signal conditioner, and an acquisition device incorporated in the chip itself. For this reason, an acquisition and processing system based on MEMS microphone arrays is reduced to two basic elements: MEMS microphone and a processing system. This processing system is usually based on an FPGA (Field Programmable Gate Array), as it is a digital system with multiple I/O nodes. The integration of the microphone preamplifier and the ADC in a single chip significantly reduces costs, if compared with solutions based on analog microphones. This is one of the main reasons why MEMS microphones are used in acoustic arrays with a large number of channels. However, cost reduction is not the only advantage of working with MEMS microphones. This technology also reduces the space occupied by the system, which makes it feasible to build arrays with hundreds or even thousands of sensors.

Arrays of MEMS microphones are specially designed for acoustic source localization [13–15]; however, they are also used in other applications such as DOA (direction of arrival) estimation for vehicles in motion [16], speech processing [17,18], turbulence measurements [19], identifying geometric dimensions and internal defects of concrete structures [20], or acoustic imaging [21–23]. The system presented in this paper has two main characteristics/novelties: (i) it is scalable using several subarrays or modules, increasing the number of sensors used, and reconfigurable in the position and orientation of the modules; and (ii) its multi-platform framework is reconfigurable. These characteristics made our system flexible and suitable for many different applications. Many of the systems found on the literature show fixed solutions to particular problems, where the hardware is not scalable or the software is not reconfigurable as in the system that is proposed in this paper. The system shown in [22] is reconfigurable, i.e., in the position of its modules, but its acquisition framework, based on an USB port, limits the maximum number of sensors that could be added to the system. For its part, the system shown by Turqueti [16] is scalable; it allows multiple arrays to be aggregated, as the system detailed in this paper. However, our system has two advantages in comparison with Turqueti's system: (i) it can be a standalone system, due to the use of a myRIO platform (embedded hardware with a FPGA and an ARM processor) as an acquisition and data processing system; and (ii) it allows for emitting sounds, making it an active system.

This paper presents in detail the framework of a system that acquires and processes acoustic images. This system is based on a multi-platform framework, where each processing task can be interchanged between the different levels of the framework. Thus, the system can be adapted to different cost and mobility scenarios by means of its reconfigurable framework. The paper also shows the evaluation of processing times of the algorithms involved in obtaining an acoustic image, including signal processing and wideband beamforming via FFT, in each subsystem of the framework. The hardware of the system is based on modules of 8 × 8 square arrays of MEMS microphones,

allowing the system to be scalable, by means of the cluster of several modules. This paper also shows the acoustic characterization of the MEMS microphones of the square array, as well as a comparative analysis of the theoretical and measured beam patterns of one of the array modules and of some clusters formed by several array modules.

Section 2 describes the hardware setup of the system and the implemented software algorithms, both defined on the basis of the requirements stated. Section 3 presents the planar array designed for the system: it introduces MEMS microphones technology and characterizes the frequency response of the microphones used, and it characterizes the array module acoustically, obtaining its beam pattern and also the beam pattern of some clusters of modules. A reconfigurable acquisition and processing framework is proposed in Section 4, and its performance is analyzed for several scenarios. Section 5 presents a case study using the system for biometric identification tasks. Finally, Section 6 contains the conclusions and future research lines.

2. System Description

In this section, the requirements for the implementation of the acquisition and processing system for a 2D array, based on MEMS sensors, are analyzed. Then, the hardware chosen for the implementation is defined, and the processing algorithms to obtain an acoustic image using beamforming techniques are explained.

2.1. Requirements

The size of an acquisition and processing system for an array depends on the number of sensors and the set of processing algorithms. Thus, it is necessary to use a very low-cost technology per channel, in order to build a viable high dimensional array. Using MEMS microphones allows for a cost reduction of two main elements of the system: sensors and acquisition systems. Digital MEMS microphones need only one digital input to be read, although the received digital data must be processed to obtain the waveform signal. Most acquisition systems are based on FPGAs, which have a large number of digital inputs, between 40 and 1400 depending on the model; so, one FPGA can acquire as much channels as the number of digital inputs it contains. Besides, FPGA processing capacity allows the system to carry out simple operations with the acquired signals without increasing costs. The processing capacity of the FPGA is not enough to obtain the final acoustic image so it is necessary that another processor with higher capacity be joined to the FPGA.

The system must be modular and scalable. Thus, for applications that require a large number of channels, it would only be necessary to add modules of lower dimension instead of designing a new system with higher capacity. In this way, reusing arrays and their processing systems reduces costs. Extra arrays, with their acquisition subsystem, can always be added in order to build higher dimensional systems. The array modules can increase the dimension of the array, but they can also form different kinds of module configurations in order to obtain specific beam patterns, improving the performance of the array, and thus of the complete system. The use of modular subsystems implies the use of a central unit that joins the data from all the modules and controls them.

Finally, the tools and programming languages to be used should also be defined. As this system has different processing platforms, one solution can be the use of a specific language for each platform, but the use of a common programming language on all platforms is desirable.

2.2. Hardware Setup

2.2.1. MEMS Array

The acoustic images acquisition system shown in this paper is based on a Uniform Planar Array (UPA) of MEMS microphones. This array, which has been entirely developed by the authors, is a square array of 64 (8 × 8) MEMS microphones that are spaced uniformly, every 2.125 cm, in a rectangular Printed Circuit Board (PCB), as shown in Figure 1. As can also be observed in Figure 1, the PCB where

the MEMS microphones are placed has square gaps between the acoustic sensors, in order to make the array as light and portable as possible.
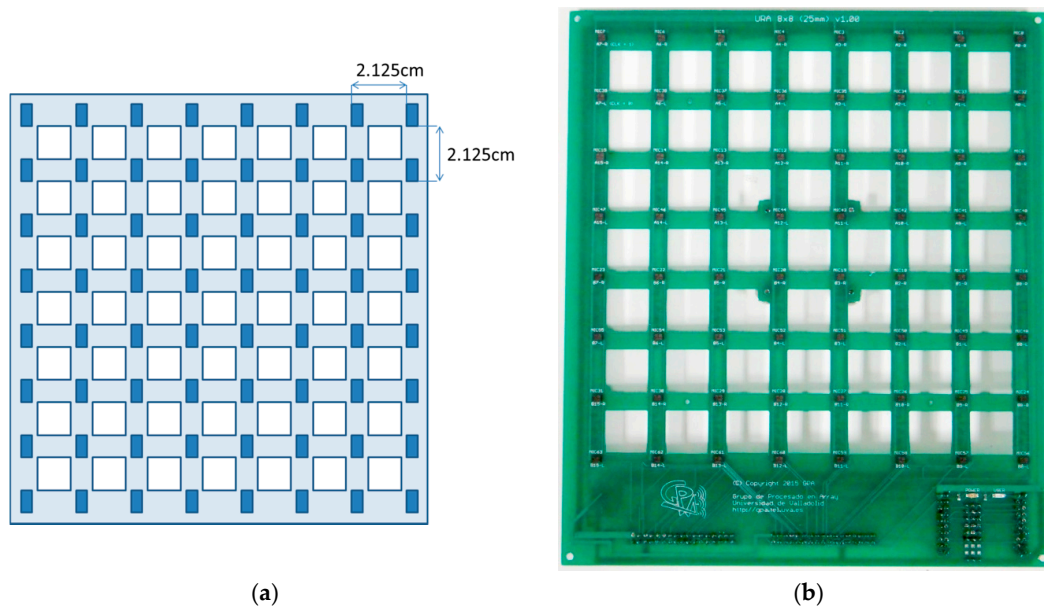

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 1. System microphone array: (**a**) sensor spacing; (**b**) system array.

This array was designed to work in an acoustic frequency range between 4 and 16 kHz. The 2.125 cm spacing corresponds to $\lambda/2$ for the 8 kHz frequency. This spacing allows a good resolution for low frequencies, while avoiding grating lobes for high frequencies in the angular exploration zone of interest [10].

For the implementation of the array, MP34DT01 microphones of STMicroelectronics were chosen. They are digital MEMS microphones with a PDM (Pulse Density Modulation) interface and with a one-bit digital signal output, obtained using a high sampling rate (1 MHz to 3.25 MHz) [24–26]. The main features of these microphones are: low-power, omnidirectional response, 63 dB SNR, high sensitivity (-26 dB FS) and an almost flat frequency response (±6 dB in the range of 20 Hz to 20 kHz).

2.2.2. Processing System

Taken into account the previous requirements, the hardware used to implement the system was selected. The design of a specific hardware was rejected due to the high cost of the design and the required time, so a search for a commercial solution was done.

MyRIO platform [27] has been selected as the base unit for this system. This platform belongs to the Reconfigurable Input-Output (RIO) family of devices from National Instruments that is oriented to sensors with nonstandard acquisition procedures, allowing low-level programming of the acquisition routines. Specifically, myRIO platform is an embedded hardware based on a Xilinx Zynq 7010 chip, which incorporates a FPGA and a dual-core ARM® Cortex™-A9 processor. The FPGA has 40 lines of digital input/output, 32 of which are used as the connection interface with the 64 MEMS microphones of the array, multiplexing two microphones in each I/O line; while the other eight lines are used to clock generation and synchronization. The ARM processor is equipped with 256 MB of DDR3 RAM, 512 MB of built-in storage space, USB Host port, and Wi-Fi interface. All this hardware is enclosed in a small box ($136 \times 89 \times 25$ mm) that costs about $1,000.

The embedded processor included in myRIO is capable of running all the software algorithms to generate acoustic images, so it can be used as a standalone array module formed by a myRIO connected to a MEMS array board as shown in Figure 2. The acoustic images can be stored in the internal storage of myRIO or in an external disk connected through the USB port.
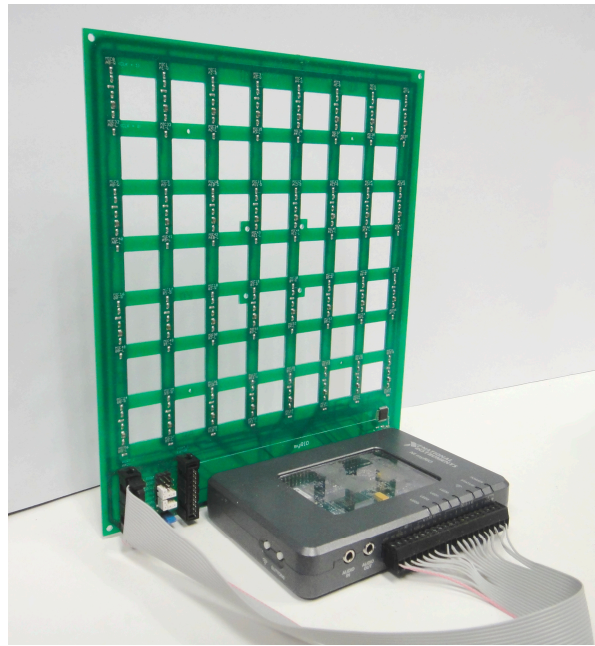
Figure 2. Array module with myRIO and MEMS array board.

Although myRIO can work as a standalone system, the lack of display means that it is usually controlled from a PC connected using a Wi-Fi interface. In a global hardware setup, as shown in Figure 3, the system includes a PC and one or more array modules. The PC performs three main functions:

- As user interface, the PC allows changing the system parameters and visualizing the acoustic images.
- As processing unit, the processors inside the PC could be used to execute the algorithms in order to obtain the acoustic images faster. Two processors are available in the PC: a general-purpose PC processor and a Graphical Processing Unit (GPU) included in the graphics card.
- As a control unit, a single PC can manage several myRIO platforms, each one associated to an array module. This feature allows clustering several modules for a proper operation of the system, which are synchronized between them using their I/O lines, where one myRIO is the master and the others are slaves.
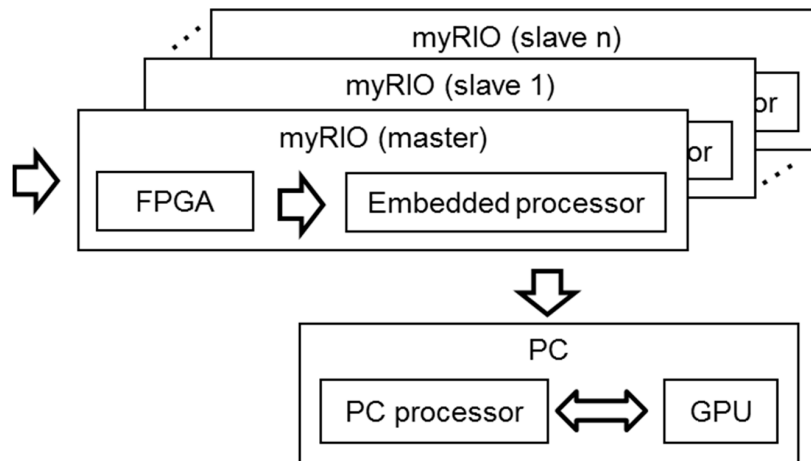

Figure 3. Global hardware setup.

2.3. Software Algorithms

The algorithms implemented in the system, shown in Figure 4, can be divided into three blocks: MEMS acquisition, signal processing, and image generation.
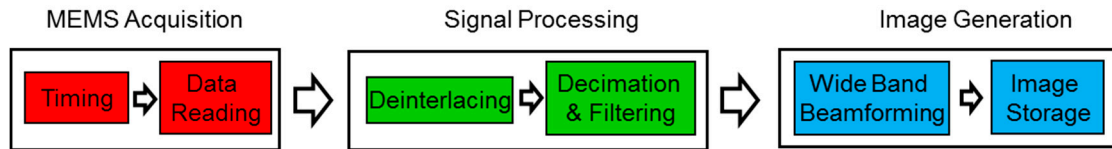

Figure 4. Software algorithms diagram.

The programming language used is LabVIEW 2015, along with its Real Time, FPGA, and GPU modules, which allows developing applications on different hardware platforms like those used in the system: FPGA, Embedded Processor (EP), PC, and GPU. In addition, most of the developed algorithms can run on any of the platforms without reprogramming.

In the acquisition block, each MEMS microphone with a PDM interface, which internally incorporates a one-bit sigma-delta converter with a sampling frequency of 2 MHz, performs signal acquisition. So, each acquired signal is coded with only one bit per sample. This block is implemented in the FPGA, generating a common clock signal for all MEMS, and reading simultaneously 64 sensors signals via the digital inputs of the FPGA. These signals are stored in 64-bit binary words, where each bit stores the signal of each MEMS. Thus, the size of the data is minimal and the transfer rate is high.

In the signal processing block two routines are implemented: (i) Deinterlacing: Through this process, 64 one-bit signals are extracted from each binary word and (ii) Decimate & Filtering: Applying downsampling techniques, based on decimation and filtering [25], 64 independent signals are obtained and the sampling frequency is reduced from 2 MHz to 50 kHz.

Finally, in the image generation block, based on wideband beamforming, a set of N × N steering directions are defined, and the beam former output are assessed for each of these steering directions. Wideband beamforming [3] computes the FFT of the MEMS signals $x_i[n]$; multiplies, element by element, each FFT $X_i[k]$ by a phase vector, that depends on the steering direction and the sensor position; and finally takes the sum of the FFT shifted in phase, as shown in Figure 5. The images generated are then displayed and stored in the system.
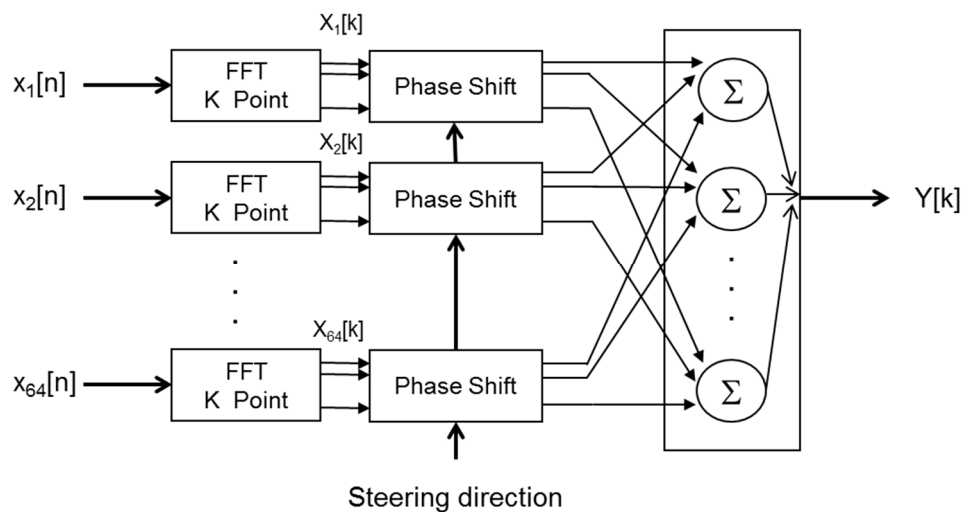

Figure 5. Wideband beamforming.

3. MEMS Array Description and Characterization

The acronym MEMS refers to mechanical systems with a dimension smaller than 1 mm [28] manufactured with tools and technology arising from the integrated circuits (ICs) field. These systems are mainly used for the miniaturization of mechanical sensors. Their small size makes interconnection with other discrete components more difficult. Therefore, when ordered, they are supplied as part of an encapsulated micro-mechanical system composed by a sensor, a signal conditioning circuit and an electric interface [29].

3.1. MEMS Microphones Characterization

An analysis of the frequency response of all MEMS microphones included in the array was performed. A sinusoidal 4 ms pulse, with a frequency changing between 2 and 18 kHz, was generated using a reference loudspeaker. Previously, the frequency response of the reference loudspeaker was calibrated using a measurement microphone (Behringer ECM 8000) and placing it in the same position as the array. Figure 6 shows the arrangement of the components used to perform this analysis. All measurements were performed in an anechoic chamber.
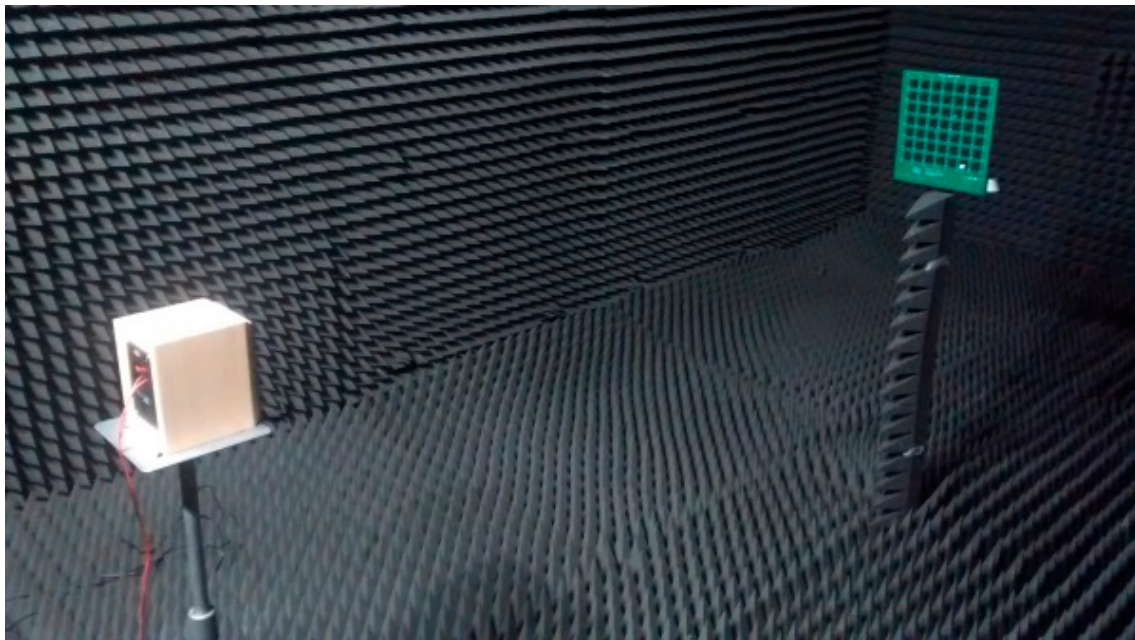

Figure 6. Arrangement of the components in the anechoic chamber.

The frequency response of each MEMS sensor was obtained and normalized according to the loudspeaker's response. Then, the average of the frequency responses was assessed. Figure 6 shows all the responses. It can be observed that the averaged frequency response is essentially flat, with a slight increase at high frequencies. This averaged response is bounded within a range of ± 3 dB. Figure 7 also shows that the frequency response of MEMS sensors varies in a range of ± 2 dB around the averaged value.
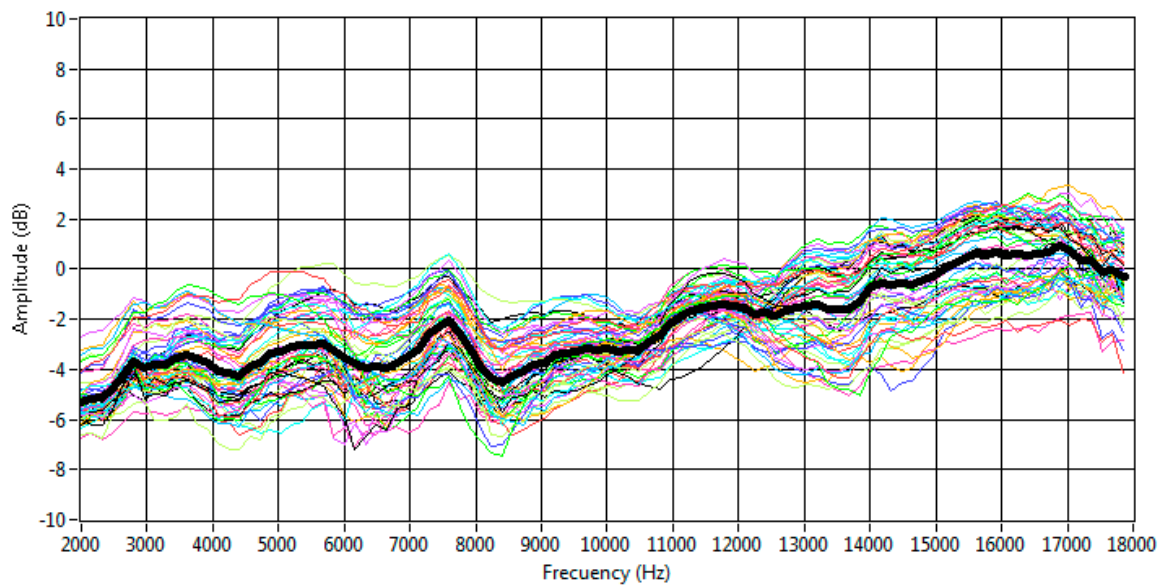
Figure 7. Frequency responses and averaged response of the MEMS microphones.

3.2. MEMS Array Characterization

3.2.1. Acoustic Characterization of an Array Module

Figure 8 shows some theoretical beam patterns of the UPA system, which are pointed towards several steering angles ([azimuth, elevation]): [$-15°$, $-15°$], [$0°$, $0°$] and [$5°$, $10°$], for several working frequencies: 4 kHz, 8 kHz and 16 kHz. It can be observed that: (i) as the frequency increases, the array angular resolution also increases, because the mainlobe beamwidth is reduced; and (ii) the sidelobe level is constant –at -13 dB for all frequencies. Figure 8b,c show that, for high frequencies, there are not grating lobes in the angular exploration zone of interest.
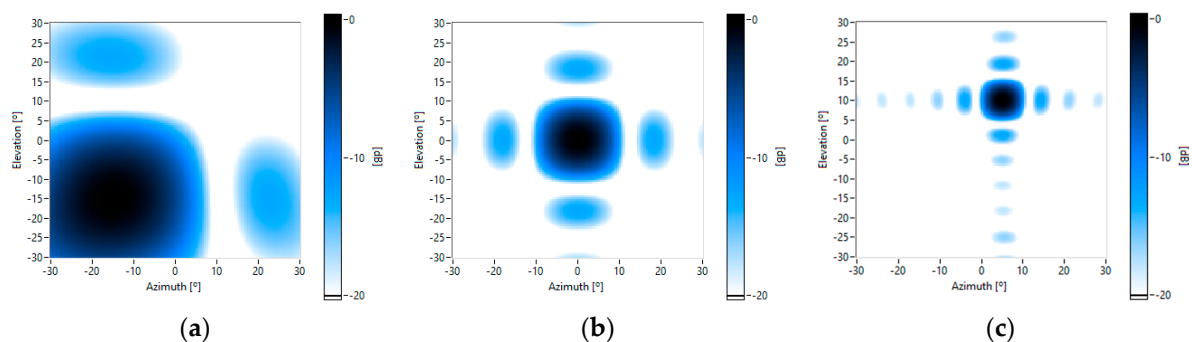

(**a**)　　　　　　　　　　　　　(**b**)　　　　　　　　　　　　　(**c**)

Figure 8. Theoretical beam patterns for different frequencies and pointing to different steering angles [azimuth, elevation]: (**a**) 4 kHz [$-15°$, $-15°$]; (**b**) 8 kHz and [$0°$, $0°$]; and (**c**) 16 kHz and [$5°$, $10°$].

For the acoustic characterization of the MEMS array, a reference loudspeaker placed in different positions was employed to obtain its beam patterns. Beamforming was carried out with a wideband FFT algorithm, focused on the loudspeaker position. Figure 9 shows some of these beam patterns.

The measured beam patterns are very similar to the theoretical ones, which assume that the acoustic sensors are omnidirectional and paired in phases. Nevertheless, a more detailed analysis of the measured beam patterns shows: (i) there are more sidelobes with a level higher than -20 dB; and (ii) there is a very small displacement of the sidelobes, which are closer. These effects are because the gain of each microphone is slightly different for each frequency, as shown in Figure 7. This is the same

effect as applying windowing techniques to the beamforming weight vector, which modifies the level and the position of the sidelobes. Thus, as the variations of the measured beam pattern, with respect to the theoretical one, are limited, it is not necessary to apply calibration techniques to the array.
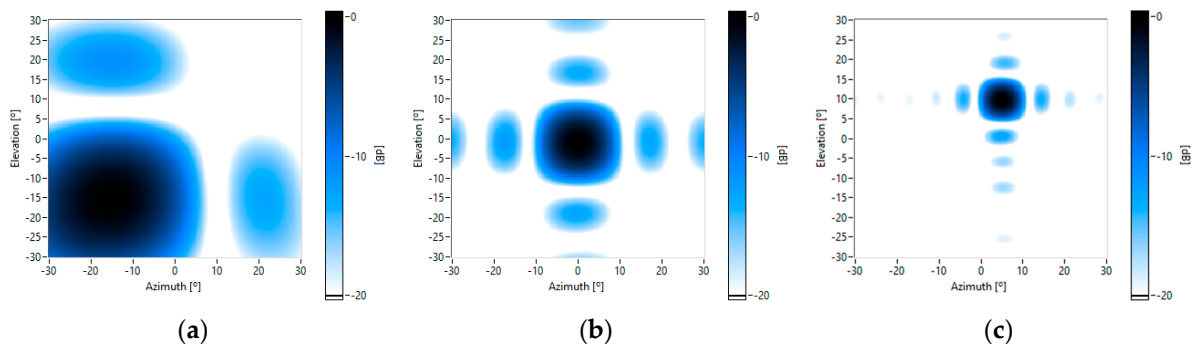

Figure 9. Measured beam patterns for different frequencies and pointing to different steering angles [azimuth, elevation]: (**a**) 4 kHz and $[-15°, -15°]$; (**b**) 8 kHz and $[0°, 0°]$; and (**c**) 16 kHz and $[5°, 10°]$.

3.2.2. Acoustic Characterization of Array Clusters

The proposed system, characterized by its modularity and scalability, can group together multiple modules with 64 sensors to build clusters with a very large number of sensors, where their geometry and spatial properties could be adapted to specific application requirements.

As an example, the acoustic characterizations of three clusters geometries are shown:

- A row cluster, to increase the directivity in one direction.
- A square cluster, to increase directivity in two orthogonal directions.
- A star cluster, to implement special radiation beam patterns.

Figure 10 shows the implemented cluster, the theoretical beam pattern, steered towards the broadside for 8 kHz, and the measured beam pattern.
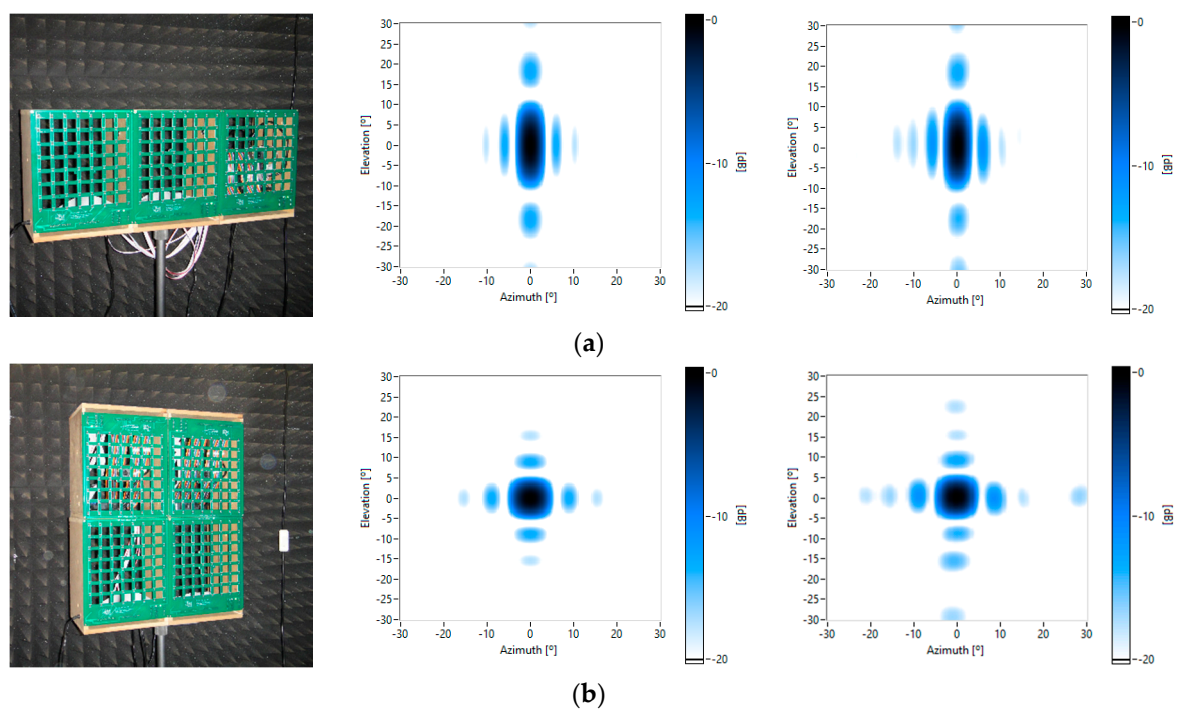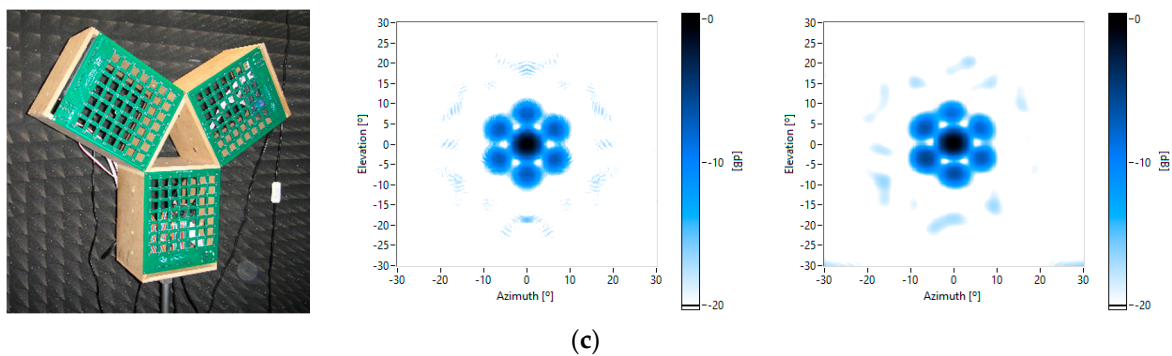

Figure 10. *Cont.*

(c)

Figure 10. Cluster (col. 1), theoretical (col. 2) and measured beampatterns (col. 3) for 8 kHz, pointed to the broadside. Cluster geometries: row (**a**); square (**b**); and star (**c**).

The row cluster shows that the beamwidth in azimuth has been reduced by a factor of 3, increasing the angular resolution of the image in that direction. In the square cluster, the beamwidth in azimuth and elevation is halved. Finally, in the star cluster, a radial symmetrical pattern is achieved with a similar beamwidth in multiple directions.

4. Multiplatform Processing Framework

On the basis of the global hardware setup presented in Section 2, a multiplatform framework with four processing levels, each one implemented over a hardware platform, was defined. These processing levels are:

- Level 1 (L1) corresponds to the FPGA based on its capacity to carry out simple tasks of filtering and decimation. The parallelization degree is maximum and limited by the number of the FPGA resources (Look Up Tables, multipliers, DSP units, RAMs, etc.).
- Level 2 (L2) is based on an Embedded Processor (EP), such as an ARM processor that picks up the FPGA signals and carries out the first processing stages. It has limited memory as well as processing and storage capacity.
- Level 3 (L3) is based on a PC processor, such as an Intel Core i5/i7 with two to four cores. It is in charge of the main processing of the application, with medium cost and consumption. It has a high processing capacity, a great amount of memory (up to 64 Gb) and storage capacity based on a disk.
- Level 4 (L4) is formed by coprocessors, which can carry out massive FFT and lineal algebra operations, such as a Graphical Processing Unit (GPU), and they have from 200 to 1200 cores.

Processing time, parallelization degree and required memory must be analyzed for all the algorithms, needed to obtain an acoustic image, described in Section 2. The objective of this analysis is to determine the platforms/levels where these algorithms can be implemented, and the optimal distribution between the available algorithms and platforms.

The time required to transfer data between levels should also be taken into account. This transfer time, in many situations, can be similar to, or even greater than the algorithm processing time. Ideally, these transfers should be minimized, in order to process data on the same level and work with a one-way processing flow, i.e., the FPGA sends data to the EP, and it sends its data to the PC. When one level is used as a coprocessor, bidirectional flows are established, i.e., between the EP and the PC processor, through a TCP-IP interface, or between the PC processor and the GPU, using a PCIe interface.

4.1. Analysis Settings

In order to analyze the performance of the algorithms in each level, a work scheme, based on an active acoustic system, was defined. This system sends a multifrequency acoustic signal that

reaches the person under test and then, the reflected signal is collected by the MEMS array. Finally, a multichannel signal is processed following the block diagram shown in Figure 11.
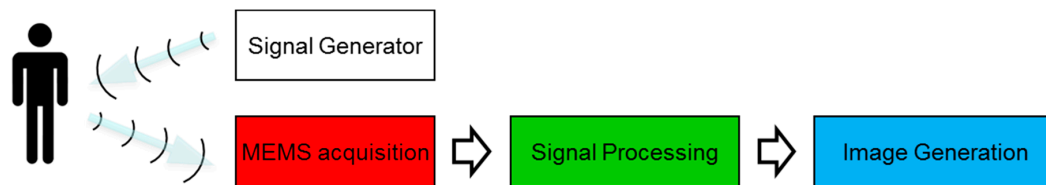

Figure 11. Processing block diagram.

There are hardware constrains that make the implementation of some algorithms in every processing level unfeasible, i.e., MEMS acquisition can only be carried out in the FPGA, or image storage cannot be executed either in the FPGA or in the GPU. Table 1 shows the main algorithms used and the processing levels where they could be implemented.

Table 1. Algorithms vs. processing level.

Algorithms	L1 FPGA	L2 EP	L3 PC	L4 GPU
MEMS acquisition	●			
Deinterlacing	●	●	●	
Decimation and filtering	●	●	●	
Wideband beamforming		●	●	●
Image storage		●	●	

The performance measurements have been carried out using the global hardware setup, with one array module, controlled by a PC. The selected PC is based on an i5 processor with four cores and 32 GB RAM, including a NVIDIA GTX 660 card with 960 cores. As algorithm parameters, an acquisition time of 30 ms, 256-point FFTs, and a grid of 40×40 steering directions have been assumed.

4.2. Performance Analysis

4.2.1. Signal Processing

The time required to implement each of the algorithms included in the signal processing on the different levels are presented in Table 2.

Table 2. Processing and transfer times [1] for the signal processing tasks.

Algorithms	L1 FPGA	L2 EP	L3 PC	L4 GPU
Processing time: Deinterlacing	0	270.3	22.4	-
Processing time: Decimation/Filtering	0	354.1	9.7	-
Transfer time: EP→PC			113.0	
Total		624.4	145.1	

[1] Time is expressed in ms.

Level 1 allows the implementation of all these algorithms using FPGA hardware resources, simultaneously with capture and signal processing without consuming additional processing time. Analyzing data from Levels 2 and 3, it can be observed that PC processing time is about 20 times lower than the time required by the EP. The times on Level 3 can be increased by transferring time from the EP to the PC for further processing. This transfer time was measured and its value is about 113 ms. Level 4, based on GPUs, was discarded because the algorithms required to perform deinterlacing and decimation/filtering are not available for this platform in LabVIEW.

4.2.2. Image Generation

Table 3 shows the required processing times related with wideband beamforming and transfer times between the PC and the GPU, for the generation of an acoustic image.

Table 3. Processing and transfer times [1] for wideband beamforming.

Algorithms	L1 FPGA	L2 EP	L3 PC	L4 GPU
Processing time: Wideband beamforming	-	257.3	18.5	25.8
Transfer time: PC→GPU→PC				7.6
Total		257.3	18.5	33.4

[1] Time is expressed in ms.

Level 1 is discarded due to the fact that (i) the FPGA included in myRIO does not have enough memory to generate and store images; and (ii) the implementation of the beamforming algorithms requires FPGAs with a large number of slices which makes it more costly. Processing times on Level 2 are the longest and the memory capacity of the EP is limited, therefore, this level is also discarded.

In order to compare Levels 3 and 4, transfer times between the PC and the GPU should be considered. The results in Table 3 show that it is preferable to perform the beamforming on the PC. These results might seem contradictory, since the processing power of the GPU is higher than the PC's. This is due to the following: (i) for 64 sensors, the GPU capacity of parallel processing is not totally used; and (ii) libraries included in LabVIEW-GPU are limited, which forces multiple data transfers between PC and GPU memories. The implementation of the beamforming algorithm in native code for the GPU and its subsequent invocation from LabVIEW would significantly reduce the overall processing time.

If the number of sensors of the system increases, by adding multiple modules of 64 sensors, the performed operations increase proportionately, taking advantage of the whole capacity of GPU parallel processing. In Table 4, the processing and transfer times versus the number of sensors is analyzed. If the number of sensors is larger than 128, GPU performance improves.

Table 4. Processing and transfer times [1] vs. number of sensors in the system.

# Sensors	PC	GPU		
	Processing Time: Wideband BF [2]	Processing Time: Wideband BF	Transfer Time: PC↔GPU	Total
64	18.5	25.8	7.6	33.4
128	49.1	40.7	8.1	48.8
256	100.2	72.9	8.7	81.6
512	198.6	138.8	11.2	150.0
1024	394.3	267.8	12.5	280.3

[1] Time is expressed in ms; [2] BF: beamforming.

Image storage can be implemented in Levels 2 and 3 because embedded microprocessors and PCs have the capacity to store data. Embedded systems incorporate a low capacity internal disk and can use external high-capacity USB drives/disks. In the case of PCs, they include both high-capacity internal and external hard drives. As storage times are very dependent of the type and model of the used media, they were discarded for the analysis.

4.3. Framework Proposals

Depending on the level where the signal processing algorithm runs, there are three implementation options. In turn, for each of these possibilities, there are up to three options depending on where the wideband beamforming algorithm is implemented. Table 5 summarizes the different options to implement the acoustic imaging algorithms with their corresponding processing and transfer times for all feasible processing levels.

Table 5. Processing and transfer times [1] for acoustic imaging algorithms.

Algorithms	Processing Level							
Processing time: Signal processing	L1: FPGA 0			L2: EP 624.4			L3: PC 32.1	
	L2: EP	**L3: PC**	**L4: GPU**	**L2: EP**	**L3: PC**	**L4: GPU**	**L3: PC**	**L4: GPU**
Wideband BF [2]	257.3	18.5	25.8	257.3	18.5	25.8	18.5	25.8
Transfer time: EP→PC	-	113.0	113.0	-	113.0	113.0	113.0	113.0
PC→GPU	-	-	7.6	-	-	7.6	-	7.6
Total	**257.3**	**131.5**	**146.4**	881.7	755.9	770.8	163.6	178.5

[1] Time is expressed in ms; [2] BF: Beamforming.

The frameworks that are based on the implementation of the signal processing algorithms through an embedded processor or a PC (grey columns) are discarded, as the FPGA allows the parallel implementation of the capture and signal processing. Thus, only three framework proposals (white columns) have been considered, associating each one to a particular use:

(1) Embedded system: The whole processing takes place in the embedded processor/FPGA without a PC, as shown in Figure 12a. This framework is optimal for applications that require portable systems and where the processing speed is not critical.

(2) PC system: The processing is shared between the FPGA and the PC. The embedded processor is used to control and transfer data between the PC and FPGA, as shown in Figure 12b. This framework presents the lowest processing time using 64 sensors. It is optimal for systems that require a short response time and/or a small number of sensors.

(3) PC system with GPU: The algorithms are implemented in the same way as in the previous framework excluding beamforming, which is implemented on the GPU, as shown in Figure 12c. This framework improves processing time as the number of sensors increases. It is the most versatile framework and it is optimal for systems that require a large number of sensors.
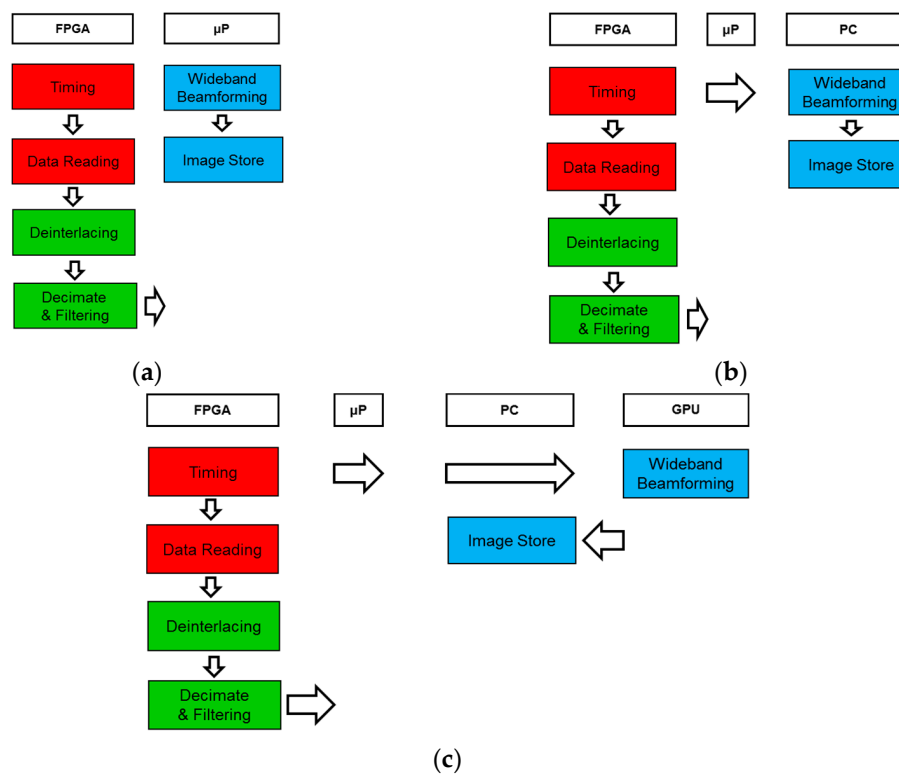
Figure 12. Embedded system (**a**); PC system (**b**); and PC system with GPU framework (**c**).

5. Case Study: Biometric Identification of People

The proposed system has multiple applications: localization and characterization of noise or vibration sources, spatial filtering and elimination of acoustic interference, etc. This case study is focused on biometric identification, as an extension of the previous system developed by the authors with a linear array of analog microphones [10]. This system could be used as an access control to enter in a medium-sized research department, where only several subjects have authorized access.

The biometric identification system is based on placing the person in front of the array and sending a multifrequency signal that is reflected on the person to be identified. The person and the system are placed inside an anechoic chamber in order to simplify the processing, but the chamber could be avoided if clutter removal techniques are used. The reflected signal is captured by the microphone array, obtaining several acoustic images for different ranges. The acoustic images are pre-processed to extract the information needed for further biometric identification. Figure 13 shows the image of a person under testing.
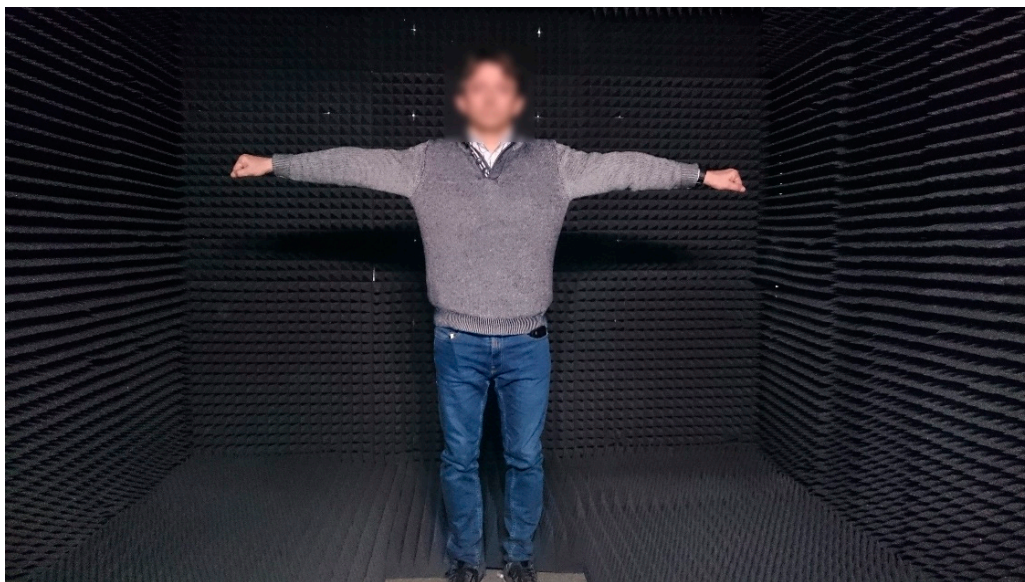

Figure 13. Example of a person under testing.

Figure 14 shows an example of some acoustic images obtained for a range of 2 m with a $\pm 22°$ angle in the azimuth coordinate and $\pm 15°$ angle in the elevation. It is observed that if the frequency increases, the spatial resolution improves and the main parts of the body could be discerned. The use of parameter extraction algorithms will improve the classification, as shown in the authors' previous work, using 1D microphone arrays [12].
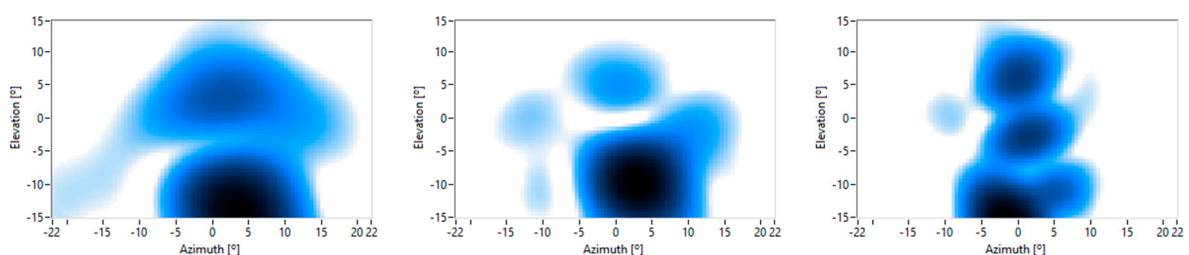

Figure 14. Acoustic images of a person with increasing frequencies, from left to right, at a constant range.

The range information can also be obtained from the captured images. Figure 15 shows images for a constant frequency and an elevation of 5°, 0° and −10°, with 0.5 m range intervals and a ±22° azimuth. Figure 15b shows that the torso is closer to the MEMS array than the arms. These results show that the designed system allows for the acquisition of 3D acoustic images.
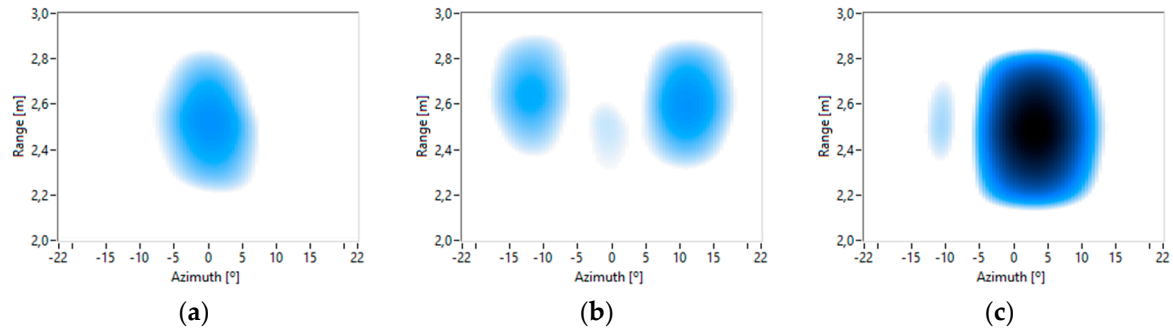

Figure 15. Acoustic images of a person at different elevations: (**a**) 5°; (**b**) 0°; and (**c**) −10° with a constant frequency.

6. Conclusions

A modular, scalable, multiplatform, and low-cost acquisition and processing system was designed to obtain acoustic 3D images. This system is based on a module with a myRIO platform and a planar array that consists of 64 MEMS microphones uniformly distributed on an 8×8 grid. The system can work with only one module or with a cluster of several of them, using the same PC unit. The system can be adapted to different cost and mobility scenarios, by means of its reconfigurable multi-platform framework, where each processing tasks can be interchanged between its different levels.

A digital MEMS microphone was selected as the acoustic sensor. This microphone allows the integration of a large number of sensors on a small sized board at low cost. An analysis of the frequency responses of these 64 MEMS microphones was carried out, obtaining: (i) a flat average response in the acoustic band, with a variation of ±3 dB; and (ii) dispersion between the responses lower than ±2 dB. The beampatterns of an array module and of different clusters of modules were also characterized, verifying that they indeed fit the theoretical models, so it is not necessary to calibrate the array.

Finally, a multiplatform framework and the necessary algorithms to obtain acoustic images from the data captured by each microphone were jointly defined. The processing time of the algorithms was evaluated on each platform. Based on these results, three different frameworks were defined for specific uses.

Thus, a versatile system for different applications, due to its modularity, scalability and reconfigurability, was designed to obtain acoustic images. Currently, the authors are using the system in the field of biometric identification, working on feature extraction and person identification from acoustic images.

Author Contributions: Juan José Villacorta and Lara Del Val Puente designed and implemented the software related with levels 0 and 1 of the framework. Alberto Izquierdo and Juan José Villacorta designed and implemented the software related with levels 2 and 3 of the framework. Luis Suárez and Juan José Villacorta designed and made the hardware prototypes. Lara Del Val Puente and Alberto Izquierdo conceived and performed the experiments, and analyzed the data. Alberto Izquierdo, Juan José Villacorta and Lara Del Val Puente wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MEMS	Micro-Electro-Mechanical Systems
FPGA	Field Programmable Gate Array
FFT	Fast Fourier Transform
ULA	Uniform Linear Array
ADC	Analog-Digital Converter
DOA	Direction of Arrival
IC	Integrated Circuit
SNR	Signal to Noise Ratio
PDM	Pulse Density Modulation
I2S	Integrated Interchip Sound
PCM	Pulse Code Modulation
UPA	Uniform Planar Array
PCB	Printed Circuit Board
GPU	Graphical Processing Unit
RIO	Reconfigurable I/O
PC	Personal Computer
OS	Operative System

References

1. Gan, W.S. *Acoustic Imaging: Techniques and Applications for Engineers*; John Wiley & Sons: New York, NY, USA, 2012.
2. Skolnik, M.I. *Introduction to RADAR Systems*, 3rd ed.; McGraw-Hill Education: New York, NY, USA, 2002.
3. Van Trees, H. *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*; John Wiley & Sons: New York, NY, USA, 2002.
4. Brandstein, M.; Ward, D. *Microphone Arrays*; Springer: New York, NY, USA, 2001.
5. Van Veen, B.D.; Buckley, K.M. Beamforming: A Versatile Approach to Spatial Filtering. *IEEE ASSP Mag.* **1988**, *5*, 4–24. [CrossRef]
6. 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. *Appl. Acoust.* **2012**, *73*, 78–82. [CrossRef]
7. Duran, J.D.; Fuente, A.I.; Calvo, J.J.V. Multisensorial modular system of monitoring and tracking with information fusion techniques and neural networks. In Proceedings of the IEEE International Carnahan Conference on Security Technology, Madrid, Spain, 5–7 October 1999; pp. 59–66.
8. Izquierdo-Fuente, A.; Villacorta-Calvo, J.; Raboso-Mateos, M.; Martínez-Arribas, A.; Rodríguez-Merino, D.; del Val-Puente, L. A human classification system for a video-acoustic detection platform. In Proceedings of the International Carnahan Conference on Security Technology, Albuquerque, NM, USA, 12–15 October 2004; pp. 145–152.
9. 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*, 10724–10737. [CrossRef] [PubMed]
10. Izquierdo-Fuente, A.; del Val-Puente, L.; Jiménez-Gómez, M.I.; Villacorta-Calvo, J. Performance evaluation of a biometric system based on acoustic images. *Sensors* **2011**, *11*, 9499–9519. [CrossRef] [PubMed]
11. Izquierdo-Fuente, A.; del Val-Puente, L.; Villacorta-Calvo, J.; Raboso-Mateos, M. Optimization of a biometric system based on acoustic images. *Sci. World J.* **2014**, *2014*. [CrossRef] [PubMed]
12. Del Val, L.; Izquierdo, A.; Villacorta, J.; Raboso, M. Acoustic Biometric System Based on Preprocessing Techniques and Support Vector Machines. *Sensors* **2015**, *15*, 14241–14260. [CrossRef] [PubMed]
13. Tiete, J.; Domínguez, F.; da Silva, B.; Segers, L.; Steenhaut, K.; Touhafi, A. SoundCompass: A Distributed MEMS Microphone Array-Based Sensor for Sound Source Localization. *Sensors* **2014**, *14*, 1918–1949. [CrossRef] [PubMed]
14. Edstrand, A.; Bahr, C.; Williams, M.; Meloy, J.; Reagan, T.; Wetzel, D.; Sheplak, M.; Cattafesta, L. An Aeroacoustic Microelectromechanical Systems (MEMS) Phased Microphone Array. In Proceedings of the 21st AIAA Aerodynamic Decelerator Systems Technology Conference and Seminar, Dublin, Ireland, 23–26 May 2011.
15. Zhang, X.; Song, E.; Huang, J.; Liu, H.; Wang, Y.; Li, B.; Yuan, X. Acoustic Source Localization via Subspace Based Method Using Small Aperture MEMS Arrays. *J. Sens.* **2014**, *2014*. [CrossRef]

16. Zhang, X.; Huang, J.; Song, E.; Liu, H.; Li, B.; Yuan, X. Design of Small MEMS Microphone Array Systems for Direction Finding of Outdoors Moving Vehicles. *Sensors* **2014**, *14*, 4384–4398. [CrossRef] [PubMed]

17. Zwyssig, E.; Faubel, F.; Renals, S.; Lincoln, M. Recognition of overlapping speech using digital MEMS microphone arrays. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 7068–7072.

18. Hafizovic, I.; Nilsen, C.; Kjølerbakken, M.; Jahr, V. Design and implementation of a MEMS microphone array system for real-time speech acquisition. *Appl. Acoust.* **2012**, *73*, 132–143. [CrossRef]

19. White, R.; Krause, J.; De Jong, R.; Holup, G.; Gallman, J.; Moeller, M. MEMS Microphone Array on a Chip for Turbulent Boundary Layer Measurements. In Proceedings of the 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Nashville, TN, USA, 9–12 January 2012.

20. Groschup, R.; Grosse, C.U. MEMS Microphone Array Sensor for Aid-Coupled Impact Echo. *Sensors* **2015**, *15*, 14932–14945. [CrossRef] [PubMed]

21. Turqueti, M.; Saniie, Y.; Oruklu, E. Scalable Acoustic Imaging Platform Using MEMS Array. In Proceedings of the 2010 IEEE International Conference on Electro/Information Technology (EIT), Normal, IL, USA, 20–22 May 2010; pp. 1–4.

22. Vanwynsberghe, C.; Marchiano, R.; Ollivier, F.; Challande, P.; Moingeon, H.; Marchal, J. Design and implementation of a multi-octave-band audio camera for real time diagnosis. *Appl. Acoust.* **2015**, *89*, 281–287. [CrossRef]

23. Sorama: Sound Solutions. Available online: www.sorama.eu/Solution/measurements (accessed on 9 June 2016).

24. Scheeper, P.R.; van der Donk, A.G.H.; Olthuis, W.; Bergveld, P. A review of silicon microphones. *Sens. Actuators A Phys.* **1994**, *44*, 1–11. [CrossRef]

25. Park, S. Principles of Sigma-Delta Modulation for Analog-to-Digital Converters. Available online: http://www.numerix-dsp.com/appsnotes/APR8-sigma-delta.pdf (accessed on 2 May 2016).

26. Norsworthy, S.; Schreier, R.; Temes, G. *Delta-Sigma Data Converters: Theory, Design, and Simulation*; Wiley-IEEE Press: New York, NY, USA, 1996.

27. Acoustic System of Array Processing Based on High-Dimensional MEMS Sensors for Biometry and Analysis of Noise and Vibration. Available online: http://sine.ni.com/cs/app/doc/p/id/cs-16913 (accessed on 14 September 2016).

28. Hsieh, C.T.; Ting, J.-M.; Yang, C.; Chung, C.K. The introduction of MEMS packaging technology. In Proceedings of the 4th International Symposium on Electronic, Materials and Packaging, Kaohsiung, Taiwan, 4–6 December 2002; pp. 300–306.

29. Beeby, S.; Ensell, G.; Kraft, M.; White, N. *MEMS Mechanical Sensors*; Artech House Publishers: Norwood, MA, USA, 2004.