PEDESTRIAN DETECTION USING AN ACTIVE MEMS MICROPHONE ARRAY ASSOCIATED TO AN AUTONOMOUS EMERGENCY BRAKING SYSTEM

Lara del Val Puente
email: lara.val@uva.es

Alberto Izquierdo Fuente
e-mail: alberto.izquierdo@uva.es

Juan José Villacorta Calvo
e-mail: juavil@tel.uva.es

Department of Signal Theory and Communications and Telematics Engineering, School of Telecommunications Engineering, University of Valladolid, Paseo del Cauce 15, 47011 Valladolid, Spain

In this work, an active sound system to detect pedestrians, based on a 2D array of MEMS microphones, is presented. It allows the detection and estimation of the position of people in environments with low visibility, solving the problems showed by the existing systems used in cars to pedestrian detection. These systems are typically based on cameras, which do not work properly in foggy or low visibility environments, and on RADARs, which detect people with low probability. This paper presents a system, working on a frequency band between 14 kHz and 21 kHz, which ensures that a vehicle travelling at up to 50 km/h can stop and avoid a collision.

Keywords: Microphone array; pedestrian detection; MEMS, Autonomous Emergency Braking

1. Introduction

Every year, more than one million road traffic collisions and accidents are associated with road vehicles, resulting in more than 300,000 pedestrian fatalities per year [1]. With the rise of autonomous cars, used in urban areas where large numbers of pedestrians are found, these figures are set to increase considerably. In an attempt to reduce this fatal toll, much work is currently underway to reduce the number of these accidents involving pedestrians.

To this effect, studies have focused on pedestrian detection and tracking [2], as well as the design of routes that avoid areas with a potential risk of pedestrian collisions [3]. Other studies focus on the development of intelligent transport systems, alerting vehicles to the presence of pedestrians in their vicinity

[4-6]. In recent times, vehicles are also being equipped with ADAS (Advanced Driver Assistance Systems) to increase safety. Among these systems are AEB (Autonomous Emergency Braking) systems, and more specifically the so-called AEB-P (AEB for Pedestrians), which focus on reducing the number of pedestrian-related accidents, for example. These systems use single-angle cameras [7], analyze time-to-collision (TTC) [8], or map the position of detected pedestrians [9].

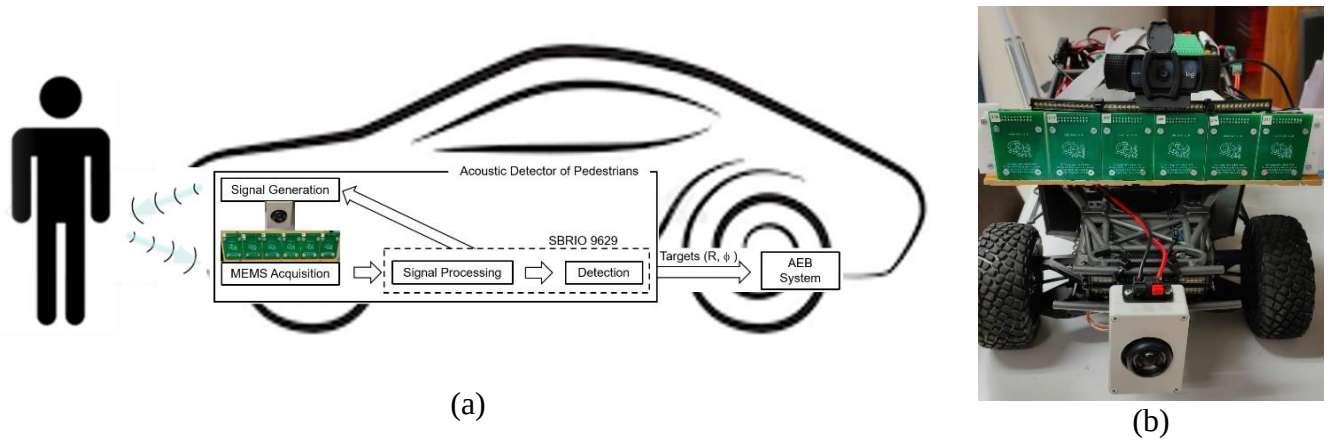
For the successful development of these studies, the use of appropriate pedestrian detection systems and algorithms is essential. Most of the systems used for pedestrian detection are based on RGB cameras and image processing algorithms, which are very effective under good visibility conditions, but perform poorly when visibility is reduced. Currently, there are several studies focused on solving this problem, either by improving pedestrian detection algorithms [10,11], or by using other sensors for pedestrian detection, information acquisition systems, such as a LIDAR [12,13], infrared sensors [14,15], or by fusing images obtained with RGB cameras with other detection systems, such as a LIDAR [16], thermal cameras [17], or a microphone array [18].

Taken this information into consideration, this paper presents a system for pedestrian detection, based on an active MEMS microphone array. The purpose of this system is to be integrated into a vehicle, together with other ADAS systems of the vehicle, to assist the AEB system in making a more reliable braking decision in a pedestrian detection in the vehicle's path. The advantage of this system over those currently in use is that acoustic signals perform well in environments with reduced visibility.

2. System description

The system presented in this article will be embedded in the front of a vehicle, in order to detect if there are pedestrians in its path, as shown in Figure 1a. As shown in Figure 1a, this system is based on the SONAR (SOund Navigation And Ranging) principle, in such a way that:

- It generates an acoustic signal.
- This acoustic signal is reflected on the possible pedestrian in the vehicle's path.
- The MEMS microphone array receives the reflected signal to be analysed and processed by the system.
- In case of detection, the system sends a warning to the vehicle's AEB system, so that it can act accordingly.



(a) (b)

Figure 1: (a) System block diagram. (b) System's acquisition system.

2.1 Hardware setup

The acoustic signal acquisition of the system is mainly based on a Uniform Planar Array (UPA) of MEMS microphones. This array is rectangular and consists of 5x30 Knowles SPH0641LU4H-1 MEMS microphones [19], as shown in Figure 1b. These sensors are distributed in 6 modules, each one with 5x5 sensors uniformly spaced every 0.9 cm. This spacing allows a good resolution of the array response for the low working frequency defined (14 kHz), while avoiding the presence of grating lobes for the high working frequencies (21 kHz).

The National Instruments sbRIO 9629 platform [20] has been chosen as the base unit of the system. This family of devices is oriented towards sensors with non-standard acquisition procedures, allowing low-level programming of the acquisition routines. In turn, the embedded processor is capable of running the software detection algorithms, so that the module formed by the sbRIO connected to the MEMS array could operate standalone. Specifically, the sbRIO 9629 used in this system is an embedded single-board controller which incorporates a FPGA Artix-7 200T and a Quad-Core Intel Atom processor.

2.2 Software algorithms

As can be observed in Figure 1a, the algorithms implemented in the system can be divided into four blocks:

- *Signal Generation block:*
  It synthesizes a pulsed multi-tone signal to be sent through the DA converter to the signal amplifier, and then to the tweeter to transmit the output signal.
- *Acquisition block:*
  Each MEMS microphone acquires the reflected signal at a 2 MHz sampling frequency. This block is implemented in the FPGA, reading simultaneously the 150 received signals via the FPGA digital inputs.
- *Signal Processing block:*
  Three routines are implemented in it:
  - A discrete set of beams are generated, covering the azimuth surveillance space for a fixed elevation angle using a Delay & Sum wideband beamformer [21]. The resolution of the system is focused on the horizontal/azimuth coordinate as it is assumed that the pedestrian will be standing on the road at the vertical height of the vehicle.
  - Using downsampling techniques (decimation and filtering), the sampling frequency of each defined beam signal is reduced from 2 MHz to 50 kHz.
  - A filter matched to the transmitted signal is then applied to the decimated signal to maximize the SNR at the input of the following block.
- *Detection block:*
  It implements three processes:
  - The relative maxima for each beam are identified and a list of potential targets is generated.
  - All targets that are outside the detection lane are eliminated, maintaining those whose distance is within the surveillance range of the system.
  - The selected targets are processed by a CFAR (Constant False Alarm Rate) detector [22]. The CFAR detector compares the energy of the potential target with a dynamic threshold that is proportional to the average of the energy received in a set of $2n_r$ cells, called reference cells, close to the evaluated one, where a number of $2n_g$ cells contiguous to the evaluated one, called guard cells, have been excluded, as can be observed in Figure 2. The CFAR threshold is obtained by varying the CFAR gain according to the equation:

$$thr_n = k \cdot \frac{1}{2n_r}\left\{\sum_{m=n-(n_r+n_g)}^{n-n_g} x_m + \sum_{m=n+n_g}^{n+(n_r+n_g)} x_m\right\} \quad (1)$$

where k is the parameter that allows weighting the relationship between the Detection Probability (Pd) and the False Alarm Probability (Pfa). The values of n_r and n_g are calculated empirically as a function of the transmitted pulse width.
In case of detection, the system sends a warning to the vehicle's AEB system, so that it can act accordingly.

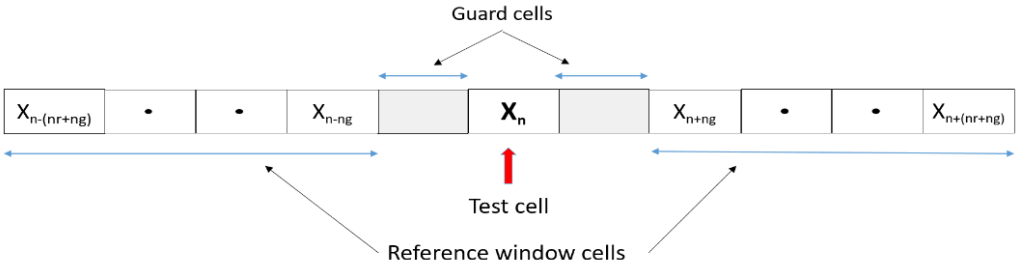
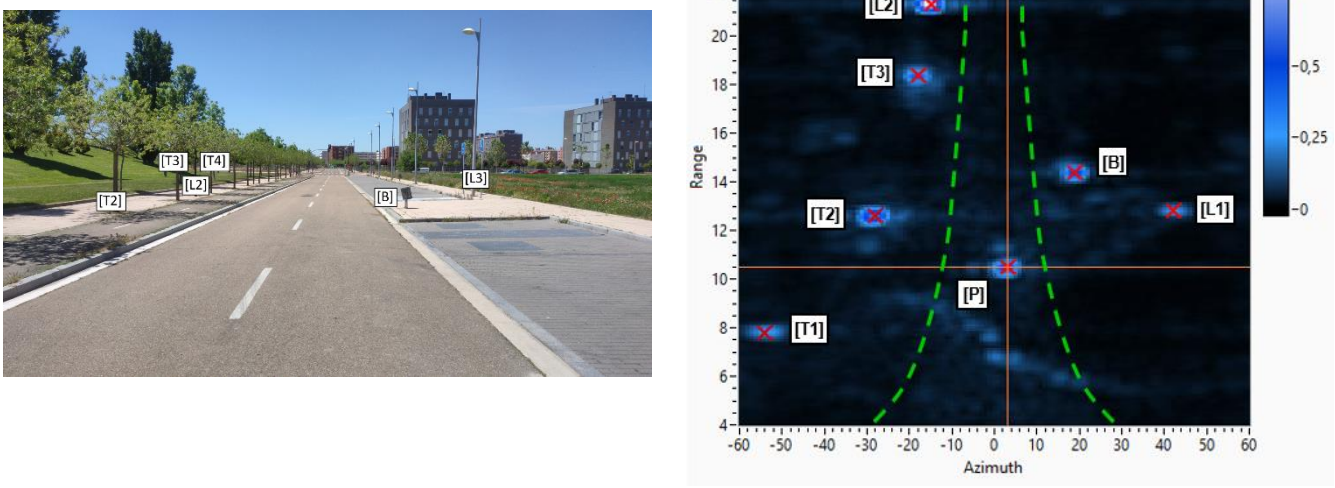


Figure 2: CFAR detector.

3. System performance and analysis

3.1 Scenario description and test setup

A stationary compact vehicle, placed on a 6 m wide road, with street lamps and trees along its edges, has been assumed for the analysis of the system performance, as it is shown in Figure 3a. Considering that the typical vehicle speed limits in urban environments are between 30 km/h and 50 km/h, and the maximum deceleration employed by its ABS braking system is of 0.8g m/s^2, the minimum braking distance to be considered speeds would be 4.4 and 12.3 m, respectively to these vehicle speeds. Taking into account these distances, 4 detection distances between the person and the acquisition system have been defined to analyse the performance of the system: 5, 10, 15 and 20 m.



(a) (b)

Figure 3: (a) Test scenario. (b) Acoustic image with detected targets for a 10 m test distance.

For a 30 km/h speed, if a pedestrian is detected at the defined test distances, the AEB system would be able to stop the vehicle, avoiding the collision with the pedestrian. However, for a 50 km/h speed, only test distances higher than 15 m would avoid the collision. In these situations, it has been confirmed, with CarSim simulation software [23], that the collision with the pedestrian would be slow enough so that the impact would not cause serious injuries.

For the tests, a signal consisting of a 3 ms multitoned pulsed signal composed of 8 1kHz-spaced tones between 14 kHz and 21 kHz, were generated with a tweeter loudspeaker. More than only one frequency has been used to improve the Detection Probability, since reflectivity on different clothes and physical characteristics of pedestrians varies with frequency [24]. The pulse width of 3 ms was chosen as a compromise between the range resolution, which is inversely proportional to the pulse width, and the transmitted energy. The SNR values that have been obtained during the tests have varied between 26 dB, for the 5 m distance, and 15 dB, for the 20 m one.

As, an example, Figure 3b show the acoustic image in azimuth/range space obtained on a test with a person located at 10 m, and for a working frequency of 20 kHz. In the acoustic image, red crosses represent the detected targets, i.e. the person (P), and other objects in front of it, such as lampposts (Lx), trees (Tx) or a bin (B) on the roadside.

In Figure 3b, it is also represented the lane boundaries by a dashed green line. It can be observed that most of the detected targets would not be considered by the system as they are outside the road lane. So, after applying the lane filter, only the pedestrian remains as a detected target.

3.2 Performance analysis

As the designed array has a 4º beamwidth ($\Delta\theta$), at 3 dB for 20 kHz, 11 equispaced beam positions has been determined to cover a ±20º azimuth area, according to the equation:

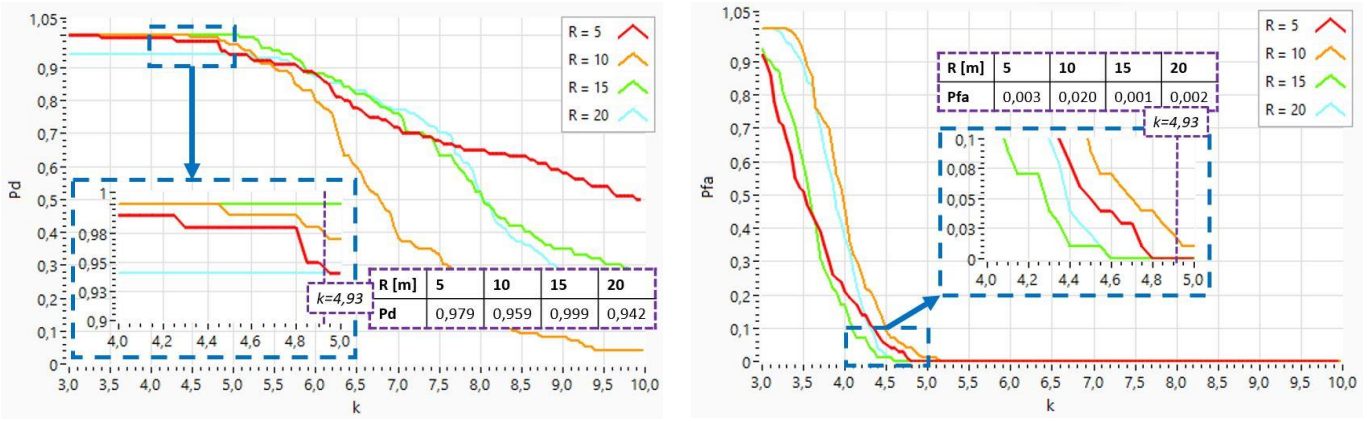
$$N_{beams} = floor\left\{2 \cdot atan2\left(\frac{d/2}{R}\right)/\Delta\theta\right\} \quad (2)$$

where a minimum range distance (R) of 5 m, and a road width (d) of 6 m have been considered.

The values of the guard and reference window cells of the CFAR detector, n_g and n_r respectively, has been obtained considering the pulse width used: 3 ms, which corresponds to a value in range of 100 cm. The suitable values found to these cells, were: a guard cell value of 200 cm and a reference cell value of 300 cm. Once these values were selected, an analysis of the CFAR detector were carried out as a function of the associated threshold gain k. This value was varied between 3 and 10, with increments of 0.01. With these operating parameters, 1000 experiments were carried out for each of the 4 defined test scenarios, with the pedestrian placed on the centre of the lane.

For each test, Figure 4 shows the Pd (Figure 4a) and the Pfa (Figure 4b) as a function of the k value, that controls the weight of the system noise power on the threshold of the CFAR detector. Furthermore, k parameter is a function of the Pfa defined for the system. Figure 4 shows that for any case both the Pd and the Pfa are monotonically decreasing functions with k. As a first approximation it is logical to think that the Pd decrease with the distance, since the SNR decreases.

Analysing the behaviour of the Pd (Figure 4a), it can be observed that, as expected, best behaviour is obtained for the 5 m distance. The behaviours of 15 and 20 m distances are very similar, and worse that the behaviour of the 5 m distance. But unexpectedly, 10 m distance has the worst behaviour. This behaviour could be due to the existence of reflections from the targets located at the road boundaries, which interfere with the reflections from the pedestrian itself. This misbehaviour is not a problem for the system, as it is not observed on the detector.



(a) (b)

Figure 4: (a) Detection Probability. (b) False Alarm Probability.

Pfa behaviour (Figure 4b) is similar to the previous one. It also reflects the influence of reflections from targets outside the lane. In this case, the behaviour can also be influenced by the noise present in the scenario (passing vehicles, acoustic signals from traffic lights to assist the blind, pedestrians passing by, etc.). This joint behaviour of Pd and Pfa is shown, in Figure 5, in the ROC curves for the different scenarios.

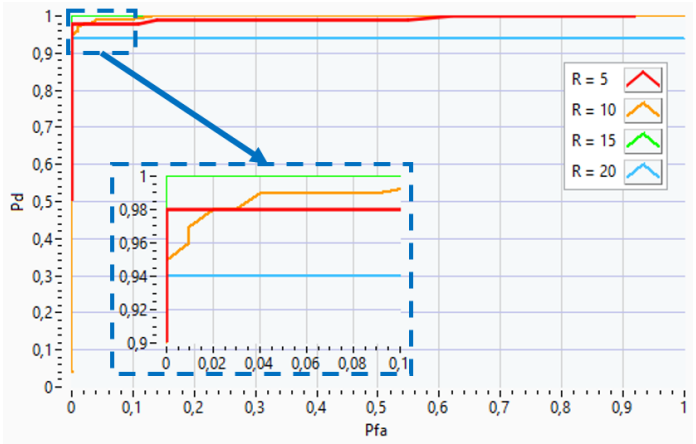


Figure 5: ROC curves for the tests scenarios.

Last step on the system configuration is the definition of the optimal detection threshold (via k), obtained from a pre-set Pfa value, according to the Neyman-Pearson Observer theorem of the RADAR theory [22]. Particularly, in this study, this Pfa value has been set to $2 \cdot 10^{-2}$, as this is a typical value used for in-vehicle detection systems.

Analysing Pfa values for each scenario (Figure 4b), it is observed that to ensure the Pfa limit value in all the scenarios, the value of k should be 4.93. For this threshold, analysing Pd values (Figure 4a), it is observed that for all the scenarios Pd values are above 0.940. These Pd values are adequate in scenarios where a person's life is at risk.

4. Conclusions

This work presents an acoustic system that estimates the position of a person up to a distance of 20 m, using a 2D array of MEMS microphones. The system is based on beamforming, range and lane filtering and finally a powerful CFAR detection algorithm.

This system will cooperate in a vehicle with an AEB system, in order to prevent the car from colliding with a pedestrian in urban environments (up to speeds of 50 km/h). For higher speeds, the AEB system, won't avoid the collision but it will mitigate the impact, avoiding fatal injuries.

The analysis of the system performance for a set of tests with a person placed at different distances has shown its feasibility. The obtained operating parameters after its analysis performance with 4000 experiments are: False Alarm Probability lower than 0.02 and Detection Probability higher than 0.94.

This system could be fused with existing camera, LIDAR and RADAR based pedestrian detection systems, in order to decrease the False Alarm Probability and increase the joint Detection Probability of the whole system. By estimating the pedestrian 2D position, the system could also warn the vehicle's steering system to avoid the pedestrian.

Acknowledgements

This research was funded by Ministerio de Ciencia, Innovación y Universidades, grant number RTI2018-095143-B-C22.

REFERENCES

1 World Health Organization. (2018). *WHO Global Status Report on Road Safety 2018: Summary. Technical Documents*. [Online.] available online: https://apps.who.int/iris/handle/10665/277370

2 Zhang, H.; Liu, Y.; Wang, C.; Fu, R.; Sun Q.; Li, Z. Research on a Pedestrian Crossing Intention Recognition Model Based on Natural Observation Data, *Sensors*, **20** (6), ID 1776, (2020).

3 Lozano-Domínguez, J.M.; Mateo-Sanguino, T.J. Walking Secure: Safe Routing Planning Algorithm and Pedestrian's Crossing Intention Detector Based on Fuzzy Logic App, *Sensors*, **21** (2), ID 529, (2021).

4 Branquinho, J.; Senna, C.; Zúquete, A. An Efficient and Secure Alert System for VANETs to Improve Crosswalks' Security in Smart Cities, *Sensors*, **20** (9), ID 2473, (2020).

5 Vourgidis, I.; Maglaras, L.; Alfakeeh, A.S; Al-Bayatti, A.H.; Ferrag, M.A. Use Of Smartphones for Ensuring Vulnerable Road User Safety through Path Prediction and Early Warning: An In-Depth Review of Capabilities, Limitations and Their Applications in Cooperative Intelligent Transport Systems, *Sensors*, **20** (4), ID 997, (2020).

6 Wang, P.; Motamedi, S.; Qi, S.; Zhou, X.; Zhang, T.; Chan, C.Y. Pedestrian interaction with automated vehicles at uncontrolled intersections, *Transportation Research Part F: Traffic Psychology and Behaviour*, **77**, 10-25, (2021).

7 Kim, B.J.; Lee, S.B. A Study on the Evaluation Method of Autonomous Emergency Vehicle Braking for Pedestrians Test Using Monocular Cameras, *Applied Sciences*, **10** (13), ID 4683, (2020).

8 Yang, W.; Zhang, X.; Lei, Q.; Cheng, X. Research on Longitudinal Active Collision Avoidance of Autonomous Emergency Braking Pedestrian System (AEB-P), *Sensors*, **19** (21), ID 4671, (2019).

9 Park, M.; Lee, S.; Kwon, C.; Kim, S. Design of Pedestrian Target Selection With Funnel Map for Pedestrian AEB System, *IEEE Transactions on Vehicular Technology*, **66** (5), 3597-3609, (2017).

10 Liu, Y.; Ma, J.; Wang, Y.; Zong, C. A Novel Algorithm for Detecting Pedestrians on Rainy Image, *Sensors*, **21** (1), ID 112, (2021).

11 Jeong, M.; Ko, B.C.; Nam, J. Early Detection of Sudden Pedestrian Crossing for Safe Driving During Summer Nights, *IEEE Transactions on Circuits and Systems for Video Technology*, **27** (6), 1368-1380, (2017).

12 Miclea, R.C.; Dughir, C.; Alexa, F.; Sandru, F.; Silea, I. Laser and LIDAR in a System for Visibility Distance Estimation in Fog Conditions, *Sensors*, **20** (21), ID 6322, (2020).

13 Goodin, C.; Carruth, D.; Doude, M.; Hudson, C. Predicting the Influence of Rain on LIDAR in ADAS, *Electronics*, **8** (1), ID 89, (2019).

14 Piniarski, K.; Pawłowski, P.; Dąbrowski, A. Tuning of Classifiers to Speed-Up Detection of Pedestrians in Infrared Images, *Sensors*, **20** (16), ID 4363, (2020).

15 Kwak, J.; Ko, B.C.; Nam, J.Y. Pedestrian Tracking Using Online Boosted Random Ferns Learning in Far-Infrared Imagery for Safe Driving at Night, *IEEE Transactions on Intelligent Transportation Systems*, **18** (1), 69-81, (2017).

16 Wei, P.; Cagle, L.; Reza, T.; Ball, J.; Gafford, J. LiDAR and Camera Detection Fusion in a Real-Time Industrial Multi-Sensor Collision Avoidance System, *Electronics*, **7** (6), ID 84, (2018).

17 Shopovska, I.; Jovanov, L.; Philips, W. Deep Visible and Thermal Image Fusion for Enhanced Pedestrian Visibility, Sensors, **19** (17), ID 3727, (2019).

18 King, E.A.; Tatoglu, A.; Iglesias, D.; Matriss, A. Audio-visual based non-line-of-sight sound source localization: A feasibility study, *Applied Acoustics*, **171**, ID 107674, (2021).

19 Knowles. (2021). *SPH0641LU4H-1 MEMS microphone*. [Online.] available: https://www.knowles.com/docs/default-source/model-downloads/sph0641lu4h-1-revb.pdf

20 National Instruments. (2021). *sbRIO-9629 Platform*. [Online.] available: https://www.ni.com/pdf/manuals/377898c_02.pdf

21 Van Veen, B.D.; Buckley, K.M. Beamforming: A Versatile Approach to Spatial Filtering. *IEEE ASSP Magazine*, **5**, 4–24, (1988).

22 Skolnik, M.I. *Introduction to RADAR systems*, McGraw-Hill Education, 3rd edition, New York, NY (2002).

23 CarSim. (2021). *Vehicle performance simulation system*. [Online.] available: https://www.carsim.com/products/carsim/index.php

24 Izquierdo, A.; Del Val, L.; Jiménez, M.I.; Villacorta, J.J. Performance Evaluation of a Biometric System Based on Acoustic Images, *Sensors*, **11** (10), 9499-9519, (2011).