A Wireless Acoustic Array System for Binaural Loudness Evaluation in Cities

Article *in* IEEE Sensors Journal · September 2017

DOI: 10.1109/JSEN.2017.2751665

CITATIONS
13

READS
103

6 authors, including:

Juan Emilio Noriega-Linares
Saint Anthony Catholic University
11 PUBLICATIONS **69** CITATIONS
SEE PROFILE

Alberto Rodriguez-Mayol
Saint Anthony Catholic University
9 PUBLICATIONS **57** CITATIONS
SEE PROFILE

Jaume Segura Garcia
University of Valencia
160 PUBLICATIONS **1,078** CITATIONS
SEE PROFILE

Santiago Felici-Castell
University of Valencia
93 PUBLICATIONS **854** CITATIONS
SEE PROFILE

A Wireless Acoustic Array System for Binaural Loudness Evaluation in Cities

Juan E. Noriega-Linares, Francisco A. Rodriguez, Maximo Cobos, *Senior Member, IEEE*, Jaume Segura-Garcia, Santiago Felici-Castel, and Juan M. Navarro

***Abstract*—Networks of acoustic sensors are being deployed in smart cities to continuously monitor noise levels. In this paper, a novel acoustic sensor device is designed for binaural loudness evaluation, in a standalone platform. The audio is acquired from an array of microphones and a binaural signal is synthesized by a Direction-Of-Arrival algorithm and a Head-Related Transfer Function. Hardware setup and software algorithms are presented and the results are discussed. Finally, the tests conducted in an early deployment show the feasibility of using the device to carry out large temporal and spatial sampling for the evaluation of binaural loudness.**

***Index Terms*—Binaural loudness monitoring, acoustic sensor network, psychoacoustic, deployment, low-cost, microphone array.**

I. Introduction

THE population of XXIst century cities is increasing so fast that new problems have appeared [1]. The problem of noise is a major concern in most modern cities, and several measurements and studies of this environmental factor have been done in the past [2]. Although, traditionally, these studies have been done using objective parameters, like the equivalent sound pressure level [3], however some research in the past have shown that studies which evaluate subjective parameters, such as loudness and sharpness, are more suitable for the assessment of noise annoyance in people [4]. Psychoacoustic research has been widely studied and standards for evaluating noise annoyance and calculating psychoacoustic parameters have been created [5], [6].

In this context, a new approach has emerged that aims to continuously monitor the sound environment of cities [7], [8], allowing the collection and analysis of huge amounts of data, thus shaping a new perspective for studying cities. Noise monitoring has helped to delve into the knowledge that people and institutions, such as city councils or regional governments, have of their cities. The use of sensor networks with an acoustic sensor has been studied in several works [9], [10], [11], [12], [13].

One of the main factors in the perceived annoyance assessment of a sound is loudness [14], [15]. Loudness is the subjective intensity of sound, an attribute that characterises sounds ranging from quiet to loud. It is primarily a psychological correlate of the amplitude of a sound. There exist some models for the calculation of this parameter which seek to process numerically an estimation of the loudness level based on the objective characteristics of the sound. The most relevant and well-known models are Zwicker's [15] and Moore's [16]. They are based on a monaural stimulus that goes through some filtering stages which correspond to the ear natural filters and the hearing system. More recently, new research in the psychoacoustics field has shown that monaural loudness is not sufficient for the acoustic loudness assessment and models for binaural loudness have been developed [17], [18].

In this paper, the design of an acoustic monitoring device that analyses the binaural loudness level from a 2-microphone array is described. The use of a mini computer, a Raspberry Pi, as the core of the device allows to perform the acquisition, the binaural synthesis and the calculation of the acoustic parameters on-board. Also, a mechanism to send the results to the cloud by an Internet-of-Things platform for later processing has been implemented. Using this online service, it is possible to perform further statistical analysis of the stored data, e.g., employing spatial statistical techniques [19].

The paper is organised as follows: in section II, the binaural synthesise from two monophonic signals recorded with two microphones is described, and the method for the transformation is exposed. After that, the binaural loudness model used in this work and the stages for developing the algorithm are explained. In section III, the description of the building of the prototype is shown. In section IV, a pilot test using a device in a particular scenario to analyse the environmental noise is described. Finally, the conclusions are presented in section V.

II. Binaural loudness evaluation process

This section describes in detail the processes involved in the binaural loudness analysis.

The section is divided into two parts. First, the 2-microphone array signal processing for synthesising the binaural signals is described, and second, the binaural loudness calculation and a review of the loudness model applied.

A. Binaural signal synthesis

Next, the signal processing approach used in this paper to synthesise a binaural signal from the output of a two-microphone array is described. The process consists of two steps. First, the microphone signals are analysed in the time-frequency domain to estimate the direction of arrival (DOA) of each time-frequency bin by assuming plane-wave propagation.

J. E. Noriega-Linares, F. A. Rodriguez and J. M. Navarro are with the Grupo de Telecomunicaciones Avanzadas, Catholic University of Murcia, Murcia, 30107, Spain (e-mail:jenoriega@ucam.edu; farodriguez@ucam.edu; jmnavarro@ucam.edu).

M. Cobos, J. Segura-Garcia and S. Felici-Castell are with the School of Technical Engineering, University of Valencia, Valencia, 46010, Spain (e-mail: maximo.cobos@uv.es; jaume.segura@uv.es; santiago.felici@uv.es).

Then, each time-frequency point is selectively filtered by the Head-Related Transfer Function (HRTF) corresponding to the estimated azimuth direction for that bin. Note that since we are using only two microphones to simplify the hardware requirements, only azimuth directions will be considered in this work. The method, which is based on the method proposed in [20], assumes sparsity of audio sources in the time-frequency domain, i.e., each time-frequency bin is assumed to belong only to one main source. As a result, the synthesised binaural signals encode signal directions in a realistic way. The method assumes that there is one major source at a given time-frequency element. Many sources can be active simultaneously, but at a given time frame their frequency overlap is assumed to be low. This assumption is known as W-Disjoint Orthogonality (WDO), and it has been shown that audio sources in most practical situations can be approximated to be WDO [21]. Secondary sources will have a different direction as long as they do not have a severe time-frequency overlap (WDO orthogonality assumption). This assumption has been already studied by the audio source separation community, and has been shown to be a good approximation for different types of sources [22], [23], [24]. Source separation algorithms based on the WDO assumption have already been successfully applied over a wide range of acoustic scenarios. In [24], the recognition of soundscape elements including noise sources has already been considered.

1) DOA Estimation: Consider the output of a two-microphone array with inter-microphone distance d in an environment where N sound sources are present. The microphone signals can be written as

$$x_m(t) = \sum_{n=1}^{N} \sum_{l=0}^{L_m-1} h_{mn}(l) s_n(t-l), \quad m = 1, 2, \qquad (1)$$

where $x_m(t)$ is the signal recorded at the mth microphone at time sample t, $s_n(t)$ is the nth source signal, $h_{mn}(t)$ is the impulse response of the acoustic path from source n to sensor m, and L_m is the maximum length of all impulse responses. The above model can also be expressed in the Short-Time Fourier transform (STFT) domain as follows

$$X_m(k,r) = \sum_{n=1}^{N} H_{mn}(k) S(k,r), \quad m = 1, 2, \qquad (2)$$

where $X_m(k,r)$ denotes the STFT of the mth microphone signal, and k and r are the frequency and time frame indices, respectively. $S_n(k,r)$ denotes the STFT of the source signals $s_n(t)$ and $H_{mn}(k)$ is the frequency response from source n to sensor m. If we assume that the sources rarely overlap at each time-frequency point, Eq.(2) can be simplified to

$$X_m(k,r) \approx H_{ma}(k) S_a(k,r), \quad m = 1, 2, \qquad (3)$$

where $S_a(k,r)$ is the dominant source at time-frequency point (k,r).

To simplify, we assume an anechoic model in which the sources are sufficiently distant to consider plane wavefront incidence. Given the small distance between the two microphones, almost every source can be assumed to be in the far-field of the array. In fact, a source can be assumed to be in the near-field when its distance to the array centre is

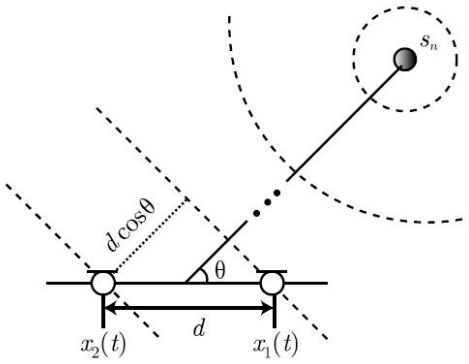

Fig. 1. Two-microphone array capturing a source signal with direction θ.

$$d < 2c(L^2)/lambda \qquad (4)$$

then,

$$d < 2(L^2)f \qquad (5)$$

Thus, in the most critical case (for a 20 Hz signal), the source should be at least at a distance greater than 16 cm, an assumption that holds in most practical situations.

Then, the frequency response is only a function of the time-delay τ_{mn} between each source and sensor, i.e., $H_{mn}(k) = e^{j2\pi f_k \tau_{mn}}$, with f_k the frequency corresponding to bin index k.

As shown in Figure 1, if the microphones are separated by a distance d and the wave from source n impinges with an angle θ, the time difference between the signals at both microphones is $(\tau_{1n} - \tau_{2n}) = \frac{d}{c}\cos(\theta)$, where c is the sound propagation speed. The phase difference observed between the two microphones in a given time-frequency bin will be

$$\Phi_{12}(k,r) = \angle\left(\frac{X_1(k,r)}{X_2(k,r)}\right) = 2\pi f_k \frac{d}{c}\cos(\theta). \qquad (6)$$

As a result, the angle corresponding to the DOA at time-frequency bin (k,r) can be estimated as

$$\hat{\theta}(k,r) = \arccos\left(\frac{c}{2\pi f_k d}\Phi_{12}(k,r)\right) \qquad (7)$$

2) HRTF Filtering: Assuming far field conditions, HRTFs are a function of the arrival direction of the source (θ) and the frequency f_k, expressed as HRTF(θ, k). Moreover, there are different HRTFs for the right and left ears, HRTF$_L(\theta, k)$ and HRTF$_R(\theta, k)$.

The synthesis strategy is simple. Any of the omnidirectional signals of the array $X_m(k,r)$ is filtered accordingly to the estimated DOA angles $\hat{\theta}(k,r)$ as follows:

$$Y_L(k,r) = X_m(k,r)\text{HRTF}_L(\hat{\theta}(k,r),k), \qquad (8)$$
$$Y_R(k,r) = X_m(k,r)\text{HRTF}_R(\hat{\theta}(k,r),k), \qquad (9)$$

where $Y_L(k,r)$ and $Y_R(k,r)$ are the STFT of the synthesised binaural signals corresponding to the left and right ears, respectively. These signals are transformed back to the time

domain using the inverse STFT operator following a classic overlap-add scheme. Note that depending on the HRTF dataset used, HRTFs are available for a set of discrete angles. It has been shown that directly using the HRTF of the available data bank closest to the estimated direction is a simple and effective approach [20]. In this paper, a bank of HRTF impulse responses with 5-degree steps is used; however, higher resolution could be used in the future, e.g., 1-degree steps, different distances or interpolation for intermediate angles.

B. Binaural Loudness Calculation

Now, the loudness calculation based on the binaural signals synthesised (Eqs. 8 and 9) is shown. These calculations are based on Zwicker's model [15] and the most recent binaural Moore's model [17]. The mentioned model attempts to process numerically the estimation of the loudness level based on the objective characteristics of the sound.

1) Monaural Loudness Model: To estimate monaural loudness levels of the synthesised binaural signals, a loudness model has been implemented. The stimuli as inputs for the model implementation are $Y_L(k,r)$ and $Y_R(k,r)$, which denote the synthesised binaural signals corresponding to the left and right ears, extracted in section II-A. Therefore, the stimulus can be defined as:

$$E_L = Y_L(k,r), \tag{10}$$

$$E_R = Y_R(k,r) \tag{11}$$

Each of the signals are passed through the different stages for the calculation of their monaural loudness level in order to, finally, extract the binaural loudness.

The processing stages that each synthesised signal has to undergo are described in Figure 2. The first and second stages of the loudness calculation correspond to the transfers through the outer and middle ear.

In [16], it is analysed the transformation from free-field sound pressure to eardrum sound pressure for a sound presented in free field from a frontal direction. The transmission function of the middle ear in the model [16] assumes roll-off at low frequencies and irregularities in the frequency range of 1.5 to 12 kHz. This is because the assumption that the excitation at absolute threshold is constant above 500 Hz.

Next, the spectral features of the signal are analysed for conversion into an excitation pattern. The excitation pattern of a sound is extracted from the spectrum reaching the cochlea. This spectrum has passed through the corrections for the effects of the outer and middle ear, and the filtered signal is the excitation pattern in a frequency spectrum of critical bands (CB). The auditory system merges sound stimuli that are close in frequency in certain critical bands. The set of these critical bands is called the critical band rate scale and it creates a frequency scale that better shows how the sound is represented in the auditory system and gives an approximation to the bandwidths of the filters in human hearing. The filters of the model follow the procedure of the ISO 532B [6]. It is measured in *Bark* [15]. This Bark scale is a frequency scale on which equal distances correspond with perceptually equal distances. Zwicker set the audible frequency range in 24 critical bands on a scale of 0 to 24 Bark.

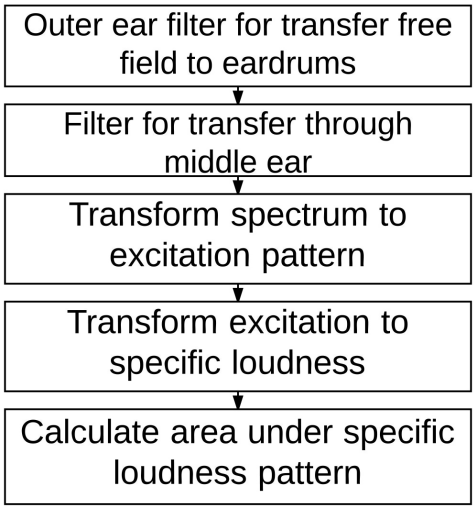

Fig. 2. Block diagram structure of the monaural loudness model procedure.

2) Monaural Loudness Calculation: Next, the transformation from an excitation produced by the sound in the auditory system to specific loudness ($N^{'}$) is carried out. The difference between the total loudness level and the specific loudness is that the latter relates to the loudness level per band, not broadband. The function that relates the excitation pattern of the stimulus, (E_{sig}) and the specific loudness, N', in a filter band is:

$$N' = 0.08 \left(\frac{E_{THQ}}{E_0}\right)^{0.23} \cdot \left[\left(0.5 + 0.5 \cdot \frac{E_{sig}}{E_{THQ}}\right)^{0.23} - 1\right] \frac{sone}{Bark} \tag{12}$$

where E_0 is the excitation that corresponds to the reference intensity $I_0 = 10^{-12} W/m^2$ and E_{THQ} is equal to the equal level loudness contour at the hearing threshold.E_{sig} corresponds to the excitation pattern of each of the signals of the stimuli (E_L and E_R).

Finally, the monaural overall loudness of the sound from each channel, N_R for the right and N_L for left, is calculated by the integration of the specific loudness per band, for all the frequency scales (Eq. 13).

$$N = \int_0^{24Bark} N'(z) \cdot dz \tag{13}$$

3) Binaural Loudness Evaluation: When a sound presented monaurally, this is perceived to be lower than when it is presented to both ears, binaurally, with the same sound level. It is the effect of binaural summation [25]. In the case where a sound is presented diotically, it would be heard twice as loud, approximately equivalent to a 10-dB increase in the sound level, as the same sound presented monaurally [26], [27]. Recent studies estimate the equivalence in around 3-8 dB [28]. For that reason, binaural loudness can provide with more information when the sound environment is analysed.

When approaching a diotic situation, the signals at the two ears tend to be weighted equally in their contribution to overall loudness [16], but the most recent studies suggest that the contributions of both ears do not follow the rule of perfect summation [17], [18]. In this paper, a model from [17]

has been implemented for the binaural loudness calculation. In Moore's binaural model, two independent calculations for both the right and left ears are performed; see section II-B1. Then, each result contributes to the overall binaural loudness. In [17], it is concluded that a signal presented diotically is approximately 1.5 times louder than the same signal presented monaurally. The value chosen is consistent with several studies presented in the mentioned work [29], [30], where the diotic signal is matched with a monaural signal in terms of loudness. Therefore, the loudness at each ear is 0.75 times the uninhibited value, giving a binaural loudness ($N_{binaural}$) of 1.5 times the monaural loudness (Eq. 14) as follows:

$$N_{binaural} = 0.75 \cdot N_R + 0.75 \cdot N_L \tag{14}$$

III. Acoustic Sensor Device

The aim of this design was to seek the conditions needed for the creation of a working prototype of an acoustic sensor device for a sensor network. The main task of this sensor is to analyse binaural loudness levels. Descriptions about the design of the device can be explained in terms of two major different parts: hardware and software. Each of them is explained in a subsection where all the different parts are broken down.

A. Design and Requirements

Some design prerequisites need to be considered, e.g., an easy installation process, to have a reliable and calibrated measuring system and the ability to analyse the sound environment and to extract manifold acoustic parameters, including the binaural loudness. The processing unit of the devices needs sufficient computing power for performing the calculations on-board and extracting all the acoustic parameters, including those that require more computational cost. The binaural loudness analysis creates the necessity of having a 2-channel sound input in order to meet the requirements of the calculation. The performance of every independent component should be good enough to carry out reliable and measurements that could span from hours to months. However, the components should also meet the premise of being low-cost in order to build affordable sensor networks of several devices with a good quality to price relation. The connectivity of the device is another fundamental pillar; therefore, the main board should provide different solutions to send and receive data.

B. Hardware

The hardware of the device includes two main parts: the audio acquisition system and the processing core. In this work, the processing core of the device is based on a Raspberry Pi [31], and the audio acquisition system consists of an array of microphones of a Sony PlayStation Eye camera. In this section, the features of both parts are described.

1) Audio Acquisition: The audio acquisition stage was based on a Sony PlayStation Eye camera with a 4-microphone array integrated into it. The decision to use this element was made following the requirements previously given in III-A. The PlayStation Eye microphone array operates with each channel processing 16-bit samples at a sampling rate of 16 kilohertz, a signal-to-noise ratio of 90 dB and a power consumption of 500 mAh. This device has been previously used in numerous studies due to the features of its camera, but in this work, the main element utilised is the microphone array for the capturing stage of the acoustic sensor device. The distance between the outermost microphones is approximately 62 milimeters, and the two middle channels are reversed. For the sound acquisition, a frequency sampling of 16,000 Hz and a bit depth of 16 bits were chosen.

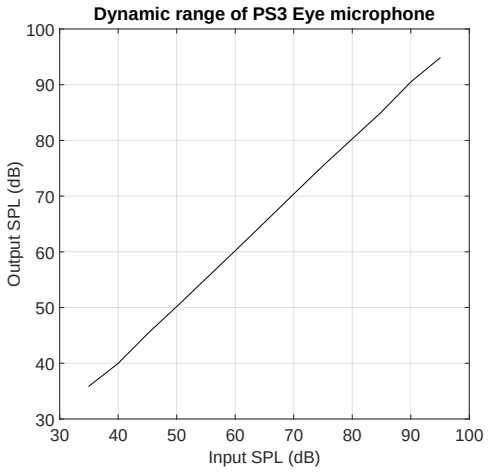

Fig. 3. Linearity and dynamic range measurements of the microphone.

To extract reliable measurements, calibration tasks were carried out with the microphone installed in the device and an integrating sound level meter, Rion NL-05 with a flat-frequency-response UC-52 microphone, in a semi-anechoic chamber using a dodecahedron speaker with a signal generator. Tests were performed to analyse the linearity and dynamic range of the microphone and to study the directivity pattern of the microphone in different frequencies. A verification using a sound level calibrator Rion NL-05 was performed before the beginning of the measurement period and after finishing it.

The audio linearity and dynamic range test shows a 60 dB dynamic range and a quite linear response of the microphone in the range of 35 dB to 95 dB (Figure 3). The directivity of the microphone was also measured showing an omnidirectional behaviour at low frequencies and a more directional behaviour as the frequency increased, as seen in Figure 4. The setting of the 2-microphone array was carried out using a Raspbian OS with the Advanced Linux Sound Architecture (ALSA). Through the laboratory tests, the microphone was adjusted in its dynamic range for being linear in a defined working range. Software updates do not affect the calibration of the signals.

2) The Processing Core: A Raspberry Pi 3 Model B computer was selected as the core of the processing, acquisition and publishing stages. The technical features of the unit include a 1.2GHz 64-bit quad-core ARMv8 CPU, 1 GB RAM, 40 GPIO pins, 4 USB ports, a full HDMI port, an Ethernet port and integrated 802.11n Wireless LAN, Bluetooth 4.1 and Bluetooth Low Energy (BLE). All these characteristics create a versatile platform, enable the development of an acoustic sensor device with the requirements stated in section III-A.

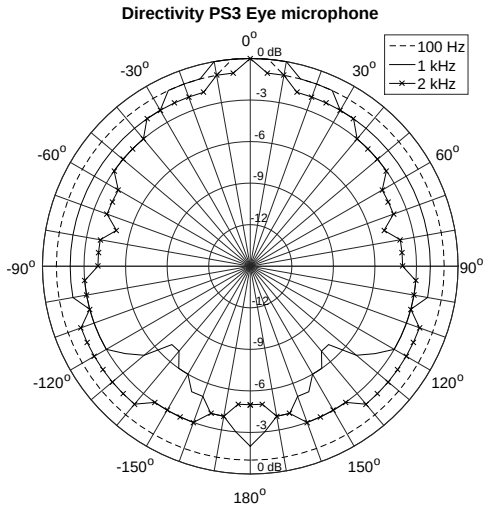

Fig. 4. Directivity pattern of the microphone for 100 Hz, 1 kHz and 2 kHz.

The USB ports and the GPIO pins are a versatile solution, providing the Raspberry Pi with a wide range of peripherals available in the market, such as WiFi antennas, ZigBee modules, microphones, cameras and connections with other devices, e.g., Arduino. A Power-Over-Ethernet connection [32] was chosen to supply the board, in which a single Ethernet cable can provide both data connection and electrical power to the device.

C. Software

The implementation of the algorithms was programmed in Python. This application consists of four main stages: audio acquisition, binaural synthesis and loudness analysis, acoustic parameter calculation and publishing of the results (Figure 5). All four processes operate independently but communicate with each other internally to ensure the well-functioning of the entire algorithm.

The algorithm consists of, first, a collection of chunks of 1 second audio from both channels, taken every 5 seconds, at a sampling frequency of 16 kHz and a 16 bit depth. The audio samples have a length of 1 second and the parameters instant sound pressure level, equivalent sound pressure level and binaural loudness are calculated for this period of time. At this point, the audio chunks take two different directions: one is to the binaural synthesis stage (see Section II-A) and the other to the acoustic parameter calculation stage. For calculating the environmental acoustic parameters, a monophonic signal is calculated using the two raw audio signals. This resulting signal will serve as input for the filtering stage and for the broadband parameter extraction. The binaural synthesis of the signals is performed, and the binaural loudness calculation is carried out, extracting the results for the monaural and binaural loudness level.

In the monophonic analysis, the signal is received and processed to extract acoustic data from the measurements. The signal takes two paths: to a filtering stage and directly to an acoustic parameter calculation stage. On the one hand, an inner filtering stage is performed in which the spectrum of the input signal is analysed. The processor passes the collected audio pieces through a set of third-octave-band filters and splits the spectrum of the sound for further sound pressure level per band calculations. This filter analyses the signal in third-octave bands from 25 Hz to 8000 Hz and extracts its sound levels in the different bands. On the other hand, the audio is processed without spectrum filtering for broadband analysis. All the calculations are performed on-board, and the values are sent to the Internet-Of-Things (IoT) platform. The chunks of audio are first processed by an A-weighted filter for the calculations of the different acoustic parameters. These parameters extracted from the audio are the instant and equivalent sound pressure level, L_p and L_{eq}, statistical noise levels L_{90}, L_{10}, L_{min} and L_{max} and third-octave band sound pressure level [33]. Parameters L_{min} and L_{max} represent the minimum and maximum equivalent levels in the measurement period. With the previous parameters obtained, the computation of hourly and daily levels is performed, obtaining the parameter L_{den} (Eq.15).

$$L_{den} = 10 \cdot lg \left(\frac{12 \cdot 10^{\frac{L_{day}}{10}} + 4 \cdot 10^{\frac{L_{evening}+5}{10}} + 8 \cdot 10^{\frac{L_{night}+10}{10}}}{24} \right) \quad (15)$$

Once all the calculations are done, the next stage of the entire process is to save and publish the results. The data obtained are stored both offline, in the local memory of the board, and online, in the cloud service, where the results are also presented. If the Internet goes down, the results can be retrieved from the internal memory of the board, serving as a backup of the online data. The SSH connection enabled on the board allows remote access for retrieving the data for extra backups to a server, if necessary.

IV. Experiment and Results

The developed unit was laboratory and field tested, to evaluate the performance of the device. Both hardware and software were assessed, and the results are presented in this section.

A deployment using a unit was installed in a residential street with a flow of people and a moderate and moving flow of vehicles for a 24-hour period. The aim of this development was to monitor the noise generated by the traffic produced by vehicles, inhabitants and services in the neighbourhood, i.e., garden care, waste treatment; and the inhabitants of the residential area. A temporal analysis will be carried out to observe the period of time with higher sound pressure levels and binaural loudness levels. A specific location on the façade of a building, at 4.5 metres height, was chosen. The location was the most appropriate in order to cover a wide variety of sound-sources and events. A cross between the two streets was at 5 meters from the façade, and the park was at 20 meters. Other residential apartments were next to the location, at 3 meters and 6 meters respectively. The environmental conditions in which the measurement was taken were: an average temperature of 27.6 degrees Celsius, a maximum temperature of 34.2 degrees Celsius, a minimum temperature of 20.9 degrees Celsius and a relative humidity of 54%.

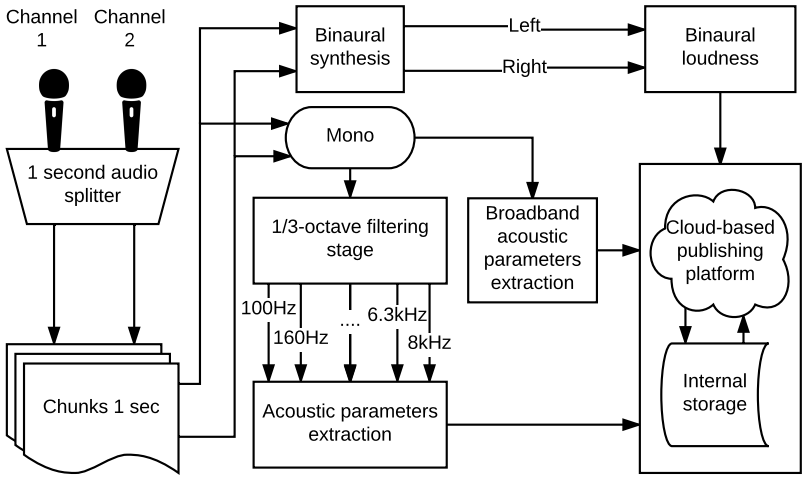

Fig. 5. Scheme of the different stages of the algorithm: audio acquisition, binaural processing and binaural loudness evaluation, parameter calculation and publishing and storing of the results.

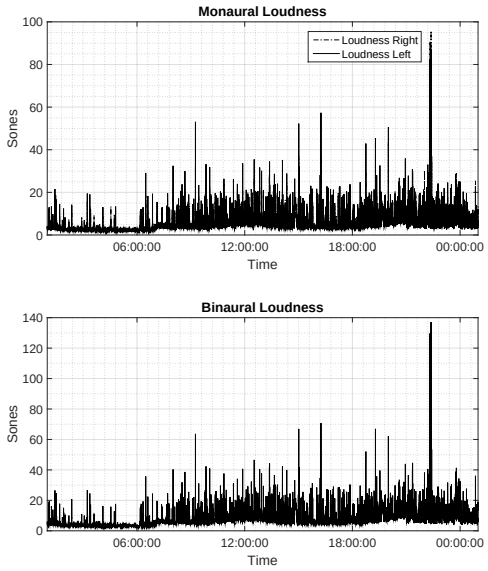

Fig. 6. Visualisation of the binaural and monaural loudness level evolution for the test period.

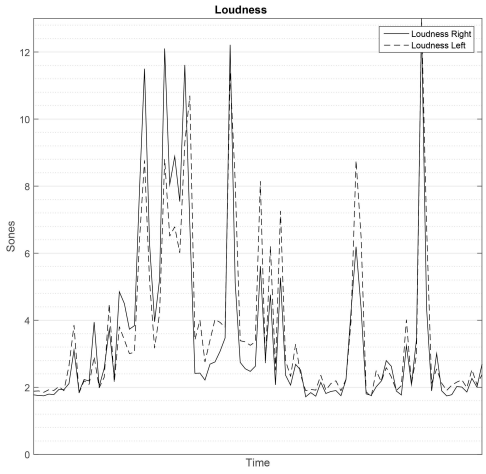

Fig. 7. Closer view of the monaural loudness levels in an event.

The loudness analysis of the period measured can be observed in Figure 6. The upper graph shows the monaural loudness levels from each of the signals acquired and synthesised, and the lower graph shows the binaural loudness level. As can be noted, they both show a similar pattern but with some small distinctions. In Figure 7, a closer view to an event shows differences between the left and right loudness levels. This is due to the direction of the event. In this case, some events with main contributions, i.e., the peaks in the graphs of loudness, from the left and some with right main contributions can be seen. A passing car from left to right or the ignition sound of a parked car to the right will show this type of behaviour.

In Figure 8, the analysis of the entire period based on sound pressure levels is illustrated. The unit performs on-board calculations of different environmental acoustic parameters, such as the L_{eq}, L_{max}, L_{min} and percentiles, as can be observed in the figure. The analysis of the equivalent sound pressure level shows differences between day and night time, distinguishing a difference of about 10-15 dB. The extraction of the parameter L_{den} is also performed in the unit calculations. When reaching 24-hour period, the unit extracts this global parameter, based on the levels extracted the rest of the day, giving an overall parameter for long-term analysis of the sound environment at that location for days, weeks or months.

In Figure 9, the equivalent sound pressure level and the binaural loudness, both in one hour period are displayed. Changes over time in both parameters shows that when one parameter increase its value, the other takes the same tendency. When the equivalent sound pressure level has low values, the binaural loudness shows low values too and in the higher levels of decibels, the loudness increases its value too. The correlation factor between the equivalent sound pressure level and the binaural loudness level in this experiment was 0.85, but as the loudness takes into account different aspects of the sound, more than its effective pressure [34], this correlation factor might not be too relevant. In Figure 10, the compar-

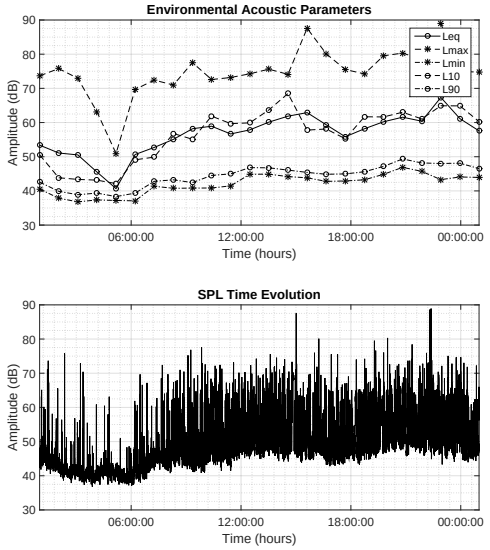

Fig. 8. Visual analysis of some environmental acoustic parameters obtained from the gathered data.

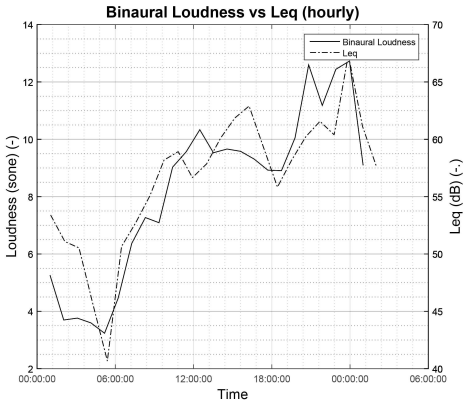

Fig. 9. Superposition of binaural loudness and equivalent sound pressure level by hour.

ison of the evolution between the binaural loudness and the monaural loudness from both channels can be observed. The correlation between monophonic and binaural loudness has been calculated.

In Figure 11, a graph with the computing processing time for the measurements of the test is shown. The efficiency of the designed unit in terms of processing time shows good performance for the needed processes and the resolution chosen for acoustical measurements. The first time stamp corresponds to recording process, and the next to the end of the binaural synthesis of the signal. After that, different marks for the different parameters are drawn. The most time-consuming process is the binaural synthesis of the signal, with an average of 2.95 seconds of processing time, while the computing of the acoustic parameters and the acquisition are performed in less than 200 miliseconds. Synthesising the binaural signal from the monophonic signal requires greater computational cost than other task of the program, due to the need to calculate the estimate of the DOA and filter with the corresponding HRTF, being the task that takes more time to be performed. However,

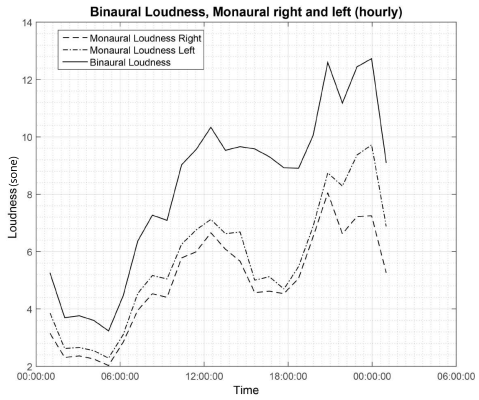

Fig. 10. Superposition of binaural loudness and monaural loudness from the right channel by hour.

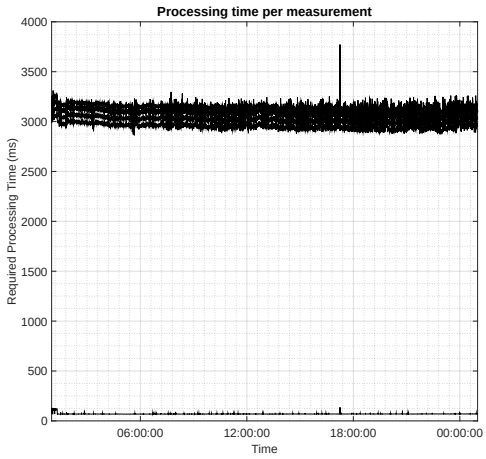

Fig. 11. Processing time for the measurements of the test for the different processes.

once the signals have been calculated, the extraction of the acoustic parameters of both signals, including the binaural loudness levels, is computed faster.

V. Conclusions

In this paper, the development of an acoustic sensor for the assessment of the binaural loudness is reported. In the past, the interest in the loudness parameter and the information that it provides to sound environment studies have been explored. However, our auditory system is binaural. Therefore, the study of the binaural loudness seemed more appropriate than that of the monaural loudness. For this reason, a two-channel acquisition system was proposed for the recording of sound. Then, a binaural synthesis process was designed and implemented for obtaining the binaural signal, i.e., a right signal and a left signal. The Raspberry Pi platform has served as an appropriate tool for the acquisition and processing of these parameters, and its connectivity has allowed remote access and publication of the results in real time. Finally, a binaural loudness model was implemented for calculations. From the experiments carried out, the suitability of performing calculations of the parameters in a standalone platform gives the flexibility of performing measurements in different places at the same time and for

periods of time that could go from days to months. With the gathering of information from the ambient noise and from the psychoacoustic parameter obtained, further analysis of these data can provide information to complement an evaluation of the acoustic environment where the units are deployed, and the information extracted can be shared with the population and the citizens.

ACKNOWLEDGMENT

The work presented in this paper has been funded by project PMAFI-02-14 by the UCAM Catholic University of Murcia, Spain, and supported by the Spanish MINECO under grant TIN2016-78799-P (AEI/FEDER, UE).

REFERENCES

[1] Loorbach, D., Shiroyama, H. The challenge of sustainable urban development and transforming cities. In Governance of urban sustainability transitions. (pp. 3–12). Springer Japan, 2016.
[2] Barrign-Morillas, J. M., Gmez-Escobar, V., Vaquero, J. M., Mndez-Sierra, J. A., Vlchez-Gmez, R. Measurement of noise pollution in Badajoz city, Spain. Acta Acust united Ac, **91(4)** (2005) 797–801.
[3] Zannin, P. H., Calixto, A., Diniz, F. B., Ferreira, J. A. A survey of urban noise annoyance in a large Brazilian city: the importance of a subjective analysis in conjunction with an objective analysis. Environmental Impact Assessment Review, **23(2)** (2003) 245–255.
[4] Fastl, H. The psychoacoustics of sound-quality evaluation. Acta Acust united Ac, **83(5)** (1997) 754–764.
[5] ISO TS 15666:2003. Acoustics - Assessment of noise annoyance by means of social and socio-acoustic surveys. 2003
[6] ISO 532B - Acoustics Method for calculating loudness level (1975)
[7] Marsal-Llacuna, M. L., Colomer-Llins, J., Melndez-Frigola, J. Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative. Technological Forecasting and Social Change (2015) **90** 611–622.
[8] Panagiotou, C. N., Zygouras, N., Katakis, I., Gunopulos, D., Zacheilas, N., Boutsis, I., Kalogeraki, V., Lynch, S., OBrien, B. Intelligent Urban Data Monitoring for Smart Cities. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML/PKDD?16), Riva del Garda, Italy, (2016).
[9] Segura-Garcia, J.; Felici-Castell, S.; Perez-Solano, J.J.; Cobos, M.; Navarro, J.M. Low-Cost Alternatives for Urban Noise Nuisance Monitoring Using Wireless Sensor Networks. *Sensors Journal, IEEE*, **2015**, *15(2)*, pp. 836–844.
[10] Santini, S.; Vitaletti, A. Wireless sensor networks for environmental noise monitoring. *Fachgesprch Sensornetzwerke*, **2007**, pp. 98–101.
[11] Farrs, J.C. Barcelona noise monitoring network. *Euronoise 2015*, **2015**, pp. 218–220.
[12] Hakala, I. Area-based environmental noise measurements with a wireless sensor network. *Euronoise 2015*, **2015**, pp. 2351–2356.
[13] Noriega-Linares, J. E.; Navarro, J. M. On the Application of the Raspberry Pi as an Advanced Acoustic Sensor Network for Noise Monitoring. *Electronics*, **2016**, *5(4)*, pp. 74.
[14] Fastl, H. Loudness and annoyance of sounds: Subjective evaluation and data from ISO 532 B. INTER-NOISE and NOISE-CON Congress and Conference Proceedings (1985) **1985(2)** pp. 1403–1406
[15] Zwicker, E., Fastl, H. (2013). Psychoacoustics: Facts and models (Vol. 22). Springer Science and Business Media.
[16] Moore, B. C., Glasberg, B. R., Baer, T. A model for the prediction of thresholds, loudness, and partial loudness. Journal of the Audio Engineering Society, (1997) **45(4)** 224–240.
[17] Moore, B. C., Glasberg, B. R. Modeling binaural loudness. The Journal of the Acoustical Society of America, (2007) **121(3)** 1604–1612.
[18] Sivonen, V. P. Directional loudness perception: the effect of sound incidence angle on loudness and the underlying binaural summation (Doctoral thesis, Acoustics, The Faculty of Engineering and Science, Aalborg University, Denmark). Retrieved from http://vbn.aau.dk/files/61782646/VPS_PhD_thesis.pdf
[19] Segura-Garcia, J.; Perez Solano, J.J.; Cobos Serrano, M.; Navarro Camba, E.A.; Felici Castell, S.; Soriano Asensi, A.; Montes Suay, F. Spatial Statistical Analysis of Urban Noise Data from a WASN Gathered by an IoT System: Application to a Small City. Appl. Sci. 2016 **6** 380.
[20] M. Cobos, S. Spors and J.J. Lopez. A Sparsity-Based Approach to 3D Binaural Sound Synthesis, in EURASIP Journal in Advances in Signal Processing, Volume 2010, Article ID 415840, 13 pages, doi:10.1155/2010/415840
[21] J. J. Burred and T. Sikora, On the Use of Auditory Representations for Sparsity-Based Sound Source Separation, in IEEE 5th International Conference on Information, Communications and Signal Processing (ICICS), 2005.
[22] J. J. Burred, "From sparse models to timbre learning: new methods for musical source separation," PhD Thesis, TU-Berlin, 2008.
[23] S. Rickard and O. Yilmaz, On the approximate W-Disjoint Orthogonality of Speech, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Orlando, Florida, 2002.
[24] O. Bunting, O. Bouzid and C. Karatsovis, "Instrument for soundscape recognition, identification and evaluation (ISRIE): technology and practical uses," in EURONOISE 2009, Edinburgh, Scotland, 2009.
[25] Reynolds, G. S., and Stevens, S. S. Binaural summation of loudness, J. Acoust. Soc. Am. 1960 **32**, pp. 1337?-1344.
[26] Fletcher, H., and Munson, W. A. Loudness, its definition, measurement and calculation, J. Acoust. Soc. Am. 1933 **5**, pp. 82-?108
[27] Stevens, S. S. 1955. The measurement of loudness, J. Acoust. Soc. Am. 1955 **27** pp. 815-?829
[28] Sivonen, V. P., Ellermeier, W. . Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summationa). The Journal of the Acoustical Society of America, 2006 **119(5)**, pp. 2965–2980
[29] Edmonds, B. A., Culling, J. F.: Interaural correlation and the binaural summation of loudness). The Journal of the Acoustical Society of America **125(6)** (2009) 3865–3870.
[30] Whilby, S., Florentine, M., Wagner, E., Marozeau, J.: Monaural and binaural loudness of 5-and 200-ms tones in normal and impaired hearinga). The Journal of the Acoustical Society of America **119(6)** 2006 3931–3939.
[31] Raspberry pi. (Online). Available: http://www.raspberrypi.org
[32] Mendelson, G. All You Need To Know About Power over Ethernet (PoE) and the IEEE 802.3af Standard. Retrieved from http://kondorsecurity.com/store/media/pdf/PoE_and_IEEE802_3af.pdf (accessed 05/07/2016).
[33] UNE-ISO 1996-1:2005. Acoustics: Description, measurement and assessment of environmental noise. Part 1: Basic quantities and assessment procedures
[34] Segura-Garcia, J., Felici-Castell, S., Perez-Solano, J. J., Cobos, M., Navarro, J. M.: Low-cost alternatives for urban noise nuisance monitoring using wireless sensor networks. IEEE Sensors Journal **15(2)** (2015) 836–844.