Guerrero, Eloi; Brugués, Josep; Verdú, Jordi; [et al.]. «Microwave microphone using a general purpose 24 GHz FMCW radar». IEEE Sensors Letters, Vol. 4, Issue 6 (June 2020), art. 3501004. DOI 10.1109/LSENS.2020.2997978

Microwave Microphone Using a General Purpose 24-GHz FMCW Radar

Eloi Guerrero, *Graduate Student Member, IEEE,* Josep Brugués, Jordi Verdú, *Senior Member, IEEE,*
and Pedro de Paco, *Senior Member, IEEE,*

***Abstract*—The usage of a commercial 24 GHz frequency-modulated continuous-wave (FMCW) radar sensor to retrieve sound signals is investigated in this article. Thanks to the great phase measurement capabilities of current radar sensors, low-frequency audio signals can be recovered by measuring the vibrations they induce in objects due to their pressure wave nature. To prove the concept, a sound emitted by a speaker is recorded by analysing the small-scale displacement of a reflective surface using a radar sensor originally devised for automotive and UAV applications at the 24 GHz ISM band.**

***Index Terms*—FMCW, microphone, radar.**

I. INTRODUCTION

RADAR sensors are now available for a wide range of applications. For example, it is possible that our vehicles feature more than four radar sensors or that traffic lights at crossroads detect the presence of pedestrians [1], [2] using microwave sensors. Thanks to integration the cost of radar systems has decreased rapidly, enabling a long list of commercial sensors to deploy.

Many previous applications have covered the use of radar sensors to retrieve small-scale vibratory motion applied to structures such as bridges or buildings aiming to monitor structural health [3], industrial machinery [4] and even loudspeakers [5]. In this scope, radar sensors compete against several alternative methods that yield different levels of accuracy. Laser Doppler Vibrometers (LDV), for example, are highly accurate but need higher maintenance and line-of-sight contact.

Among vibrations, detection may focus in those generated by sound waves. Since a pressure wave impinging into a body will trigger very small vibrations over the object's surface, detecting them enables turning sensors into microphones. This approach has been investigated by Davis et al. in [6] using high-speed video information of an object. They propose the so-called visual microphone. However, the use of cameras is limited to direct line-of-sight with well illuminated objects and need external sources of light. Moreover, sound-triggered vibrations have been previously used for remote sound eavesdropping in terms of surveillance and security. From a low-power millimetre-wave perspective, previous approaches have focused on the detection of sound, as in [7], showing how to detect heart sounds caused by muscular contraction using radar systems, or more specifically, the works in [8], [9], aiming to retrieve sound by pointing a radar sensor to the vocal cords of a human being at the 24 GHz and 94 GHz bands, respectively. Also [10] explores voice recognition from micro Doppler data introducing neural networks to classify letters and vowels. However, in all cases mentioned above, the sensors were continous-wave (CW) radars making use of the well-known CW Doppler approach.

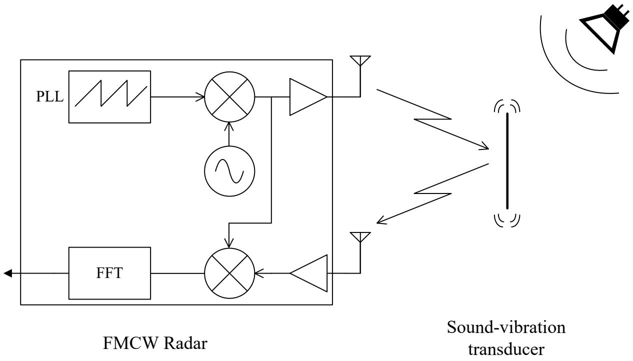


Fig. 1: FMCW sound retrieval scheme.

In this paper, we approach the measurement of sound by means of a frequency-modulated continuous-wave (FMCW) radar at the 24 GHz ISM band that is specifically designed for automotive and UAV applications, thus proving the application on a general purpose low-cost sensor. This is possible thanks to the great phase measurement capabilities that the FMCW scheme provides.

Fig. 1 depicts the structure of the proposed microwave microphone. It is composed of a phase-locked loop (PLL) generating the chirp signal, a transmitter chain for amplification, a transmitting antenna, one (or more) receiving antenna, a receiver chain to down-convert the received signals using the reference chirps and an intermediate frequency (IF) output for further processing.

II. WORKING PRINCIPLE

As opposed to the continuous-wave approach to radar in which a single unmodulated tone is emitted, the FMCW approach consists in the emission of a train of N_p chirp pulses of a certain bandwidth. As introduced in [11], information on the range of present targets can be extracted by mixing the received signals with the reference train of chirps. A received chirp will suffer from a small delay (t_{delay}) that will translate to an IF tone through the mixing process. The IF signal can be decomposed into its spectral components by means of the Fast

---

This work was supported by MCIU-Spain under Grant RTI2018-096019-B-C33.

Corresponding author: Eloi Guerrero (e-mail: eloi.guerrero@uab.cat).

Fourier Transform (FFT) on a signal processing module. Using (1), the frequency components can be translated to range in metres, being f_{IF} the IF spectral component, T_s the up-ramp time of the chirp pulse, B the swept bandwidth on each chirp and c_0 the velocity of light in vacuum.

$$R = \frac{c_0 \; f_{IF} \; T_s}{2B} \tag{1}$$

Considering (1) and the properties of FFT, targets will be resolved as different components in the spectrum if separated at least $1/T_s$. Thus, the range resolution can be defined as in (2).

$$\Delta R = \frac{c_0}{2B} \tag{2}$$

At the 24 GHz ISM band, the available bandwidth is 250 MHz, that is an achievable range resolution of 0.6 metres. Therefore, the sub-millimetre mechanical oscillations caused by sound waves cannot be extracted from the range information. Consider now a moving target. Not only time delay with respect to the reference chirp will appear but in fact, a frequency shift due to Doppler effect will be observed. Thus, the signal at the output of the receiver mixing stage will be defined as in (3), being f_c the carrier frequency and v the radial velocity of the target with respect to the radar. The expression has two unknowns: range and velocity.

$$f_{IF} = \frac{B \; t_{delay}}{T_s} \pm \frac{2 f_c v}{c_0} \tag{3}$$

The output of the initial range FFT is the magnitude and phase at each of the frequency bins (range bins using (1)). If the target suffers a small displacement between two consecutive chirps, the output of their range FFT will be equal in magnitude but different in phase. For an unambiguous detection of this displacement, it must fall into the $\pm\lambda/4$ interval between chirps, avoiding cycle slips. Then, if a frame of N_p consecutive chirps is emitted and received, the output of the range FFT of each chirp can be stacked column-wise. For a given range bin, the phase evolution can be analysed along chirps to obtain the velocity of the targets present in that bin. This is possible by computing another FFT but now in row direction; the so-called 2D-FFT that yields the well-known Range-Doppler map.

Clearly, a relation between a small displacement and a variation in phase can be established as follows:

$$\Delta\phi = \frac{4\pi\Delta r}{\lambda} \tag{4}$$

In the case of a vibratory motion, consider an object that vibrates due to a mechanical effort. The object will oscillate with a given amplitude and frequency. In most cases, this displacement is almost imperceptible to the naked eye, in the range of tens of microns, but can be measured by inspection of the phase evolution along chirps that is detected by the radar sensor [12]. From the phase, vibration frequency and amplitude can be recovered. What we propose is to use the radar to measure the vibrations due to sound waves of a reflective object and then apply a spectral estimation algorithm to recover the sound that caused the vibrations. In general terms, the algorithm in use is shown in Fig. 2.

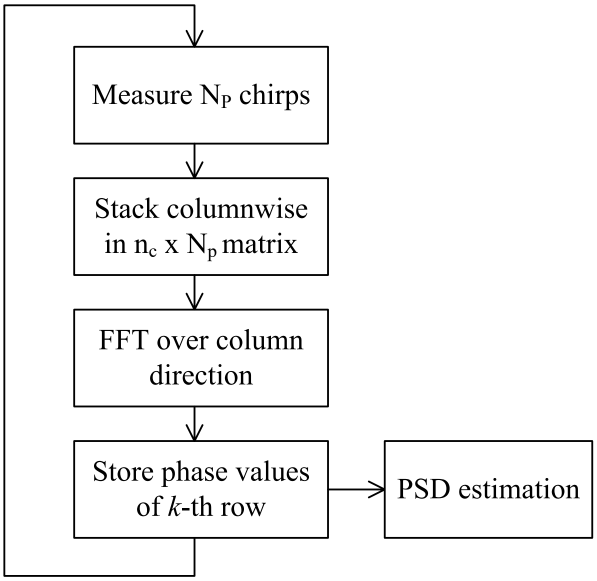


Fig. 2: Flow diagram of the proposed FMCW sound retrieval method. n_c is the number of samples per chirp and k is the range position at which the reflective object is found.

After receiving N_p chirps, each resulting IF tone is transformed to a range profile and the position of the vibrating object is found. Now, the phase value at that range position along the N_p chirps is stored in an array and a spectral estimation algorithm such as Welch [13] is applied to obtain the frequency components of the sound that induced the vibration. This process can be applied over many frames for a continuous measurement. Since FMCW radar sensors measure N_p chirps before coherently processing them, a small idle time between frames is necessary and implies a loss of phase coherence when performing a continuous measurement of sound. However, our sensor has a total chirp time of 284 μs and its idle time is in the order of 500 μs. Such idle time does not imply a loss of sound information but the appearance of low amplitude components at higher frequencies due to the loss of coherence. These components can be filtered by further processing of the data.

III. EXPERIMENTAL VALIDATION

The experimental validation of this article is based on the 24 GHz radar platform DemoRad by Analog Devices and Inras. This is a development platform with 2 transmitter and 4 receiving antennas that features an Analog Devices RFIC family designed for automotive and UAV applications. The measurement set-up is mounted inside an RF anechoic chamber. A 23x13 cm copper plate of 0.5 mm thickness is attached to one of the chamber poles and a small speaker cone is placed at a distance of 5 cm from the plate. The speaker is connected to a function generator to produce the sound that will induce vibration on the metal plate. At a distance of 2.5 meters, at the other chamber pole, the radar is mounted facing the plate. The experimental set-up is shown in Fig. 3.

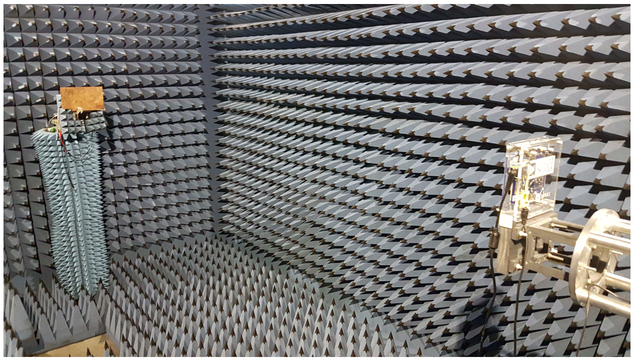


Fig. 3: Measurement set-up in the anechoic chamber.

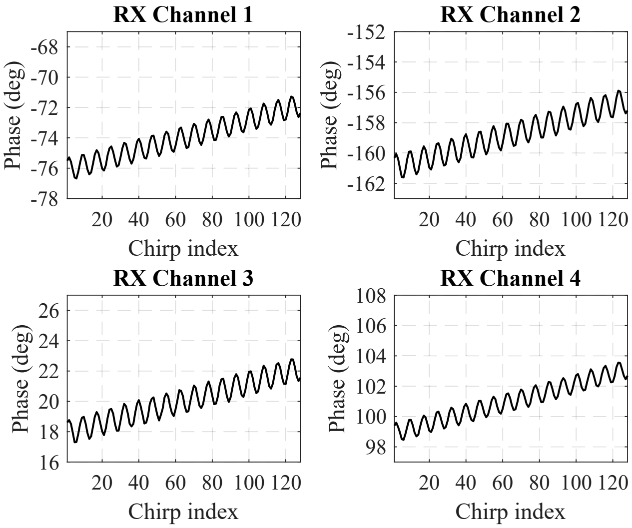
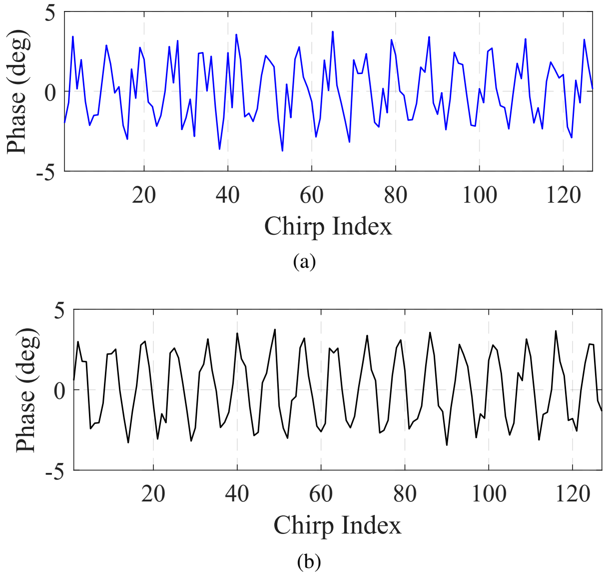


Fig. 4: Phase evolution for a 440 Hz vibration with phase drift on all four receiving channels of the DemoRad platform.



(a)

(b)

Fig. 5: Phase oscillation for an excitation of 440 Hz. (a) direct LoS, (b) behind EPS cover.

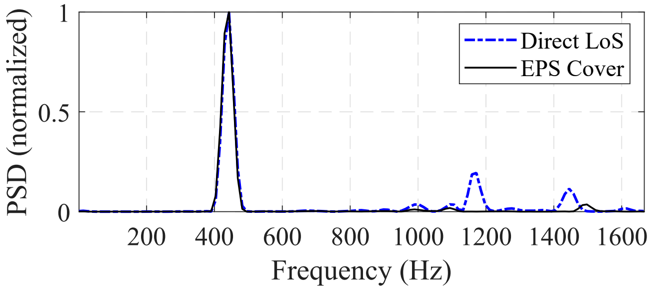


Fig. 6: Power spectral density of the sound retrieved by the radar for a 440 Hz sound excitation.

An initial simple test is carried out using as single frequency tone of 440 Hz at the speaker. A frame of $N_p = 128$ chirps is measured and the result of analysing the phase at the range bin where the plate is resolved is shown in Fig. 4. Clearly, the vibration can be observed but superposed to an undesired phase drift. This drift had been previously observed and reported to the manufacturer. This is a common issue in many implementations due to slight PLL clock drifts. In the scope of velocity detection is not a considerable issue since it implies a velocity error of ± 0.006 m/s, well below the velocity resolution of the radar. However, when trying to detect small scale vibrations the drift must be avoided, to reduce its effect on the estimation of the frequency.

To avoid the phase drift along chirps our procedure implies differentiating the phase evolution with respect to time at expense of loosing one of the 128 samples. With this, the phase result of a single receiver channel is now shown in Fig. 5a and 5b after differentiation. This figure depicts two measurements, one with line of sight between radar and plate and the other adding a wall of expanded polystyrene (EPS) between the two. This is just a simple proof of the capability of detecting sound without the need of visual contact. Notice that the phase excursion is of $\pm 3.5°$, in turn a displacement of ± 0.121 mm.

Having computed the phase evolution along chirps, thanks to the Welch spectral estimation method, the sound spectrum can be computed as shown in Fig. 6. Note that due to the differentiation stage only $N_p - 1$ samples of the phase are processed. The loss of a sample has an negligible effect in the spectral estimation resolution in this case.

Once the concept is proven for a single frame, the experiment is now to analyse how does the radar recover a sound that changes over time, as would be a conversation or a song. To do this the analysis takes into account more than a single frame of chirps. To test these continuous time approach the set-up is the same but now configuring the function generator to change from 440 Hz to 550 Hz and back within a time of 0.3 seconds. After differentiation, the spectral density can be computed another time and is plotted with respect to time, creating the spectrogram in Fig. 7.

The validation shown in this paper is performed using a single receiver channel. In a more realistic scenario related to automotive radar, it is possible to separate sound information from different vibration sources applying MIMO techniques thanks to the multiple transmit and receive channels common in this platforms. Concerning the spectrum occupation of the

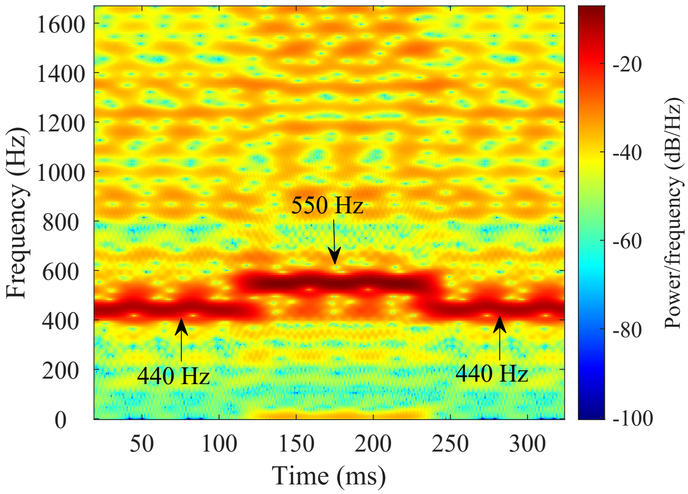


Fig. 7: Sound spectrum evolution for the 440-550-440 Hz series.

24 GHz ISM band, it is possible to mitigate interference from other automotive radar sensors applying de-ramping techniques as shown in [14].

IV. CONCLUSION

In this paper the use of a commercial FMCW radar to recover sound has been explored and demonstrated. Previous approaches in the literature used full continuous-wave custom radars to measure sound. In this case, the chirp rate used in automotive and UAV applications can achieve a detectable audible range of 3.4 kHz, that of the telephonic line, what indicates that this application can recover voice-like quality sound. The experiments also demonstrate the capability of easily detecting a displacement in the order of hundreds of microns.

REFERENCES

[1] R. Pérez, F. Schubert, R. Rasshofer, and E. Biebl, “Single-frame vulnerable road users classification with a 77 ghz fmcw radar sensor and a convolutional neural network.” IEEE, 2018, pp. 1–10.

[2] A. Bartsch, F. Fitzek, and R. H. Rasshofer, “Pedestrian recognition using automotive radar sensors,” *Advances in Radio Science*, vol. 10, pp. 45–55, sep 2012.

[3] M. Pieraccini and L. Miccinesi, “An interferometric MIMO radar for bridge monitoring,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 9, pp. 1383–1387, sep 2019.

[4] G. Vinci, S. Lindner, S. Mann, F. Barbon, S. Linz, R. Weigel, and A. Koelpin, “Six-port microwave interferometer radar for mechanical vibration analysis,” in *2013 European Radar Conference*, 2013, pp. 287–290.

[5] A. Izzo, L. Ausiello, C. Clemente, and J. J. Soraghan, “Loudspeaker analysis: A radar based approach,” *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1223–1237, feb 2020.

[6] A. Davis, M. Rubinstein, N. Wadhwa, G. Mysore, F. Durand, and W. T. Freeman, “The visual microphone: Passive recovery of sound from video,” *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 33, no. 4, pp. 79:1–79:10, 2014.

[7] C. Will, K. Shi, S. Schellenberger, T. Steigleder, F. Michler, J. Fuchs, R. Weigel, C. Ostgathe, and A. Koelpin, “Radar-based heart sound detection,” *Scientific Reports*, vol. 8, no. 11551, 2018.

[8] Heng Zhao, Zhengyu Peng, Hong Hong, Xiaohua Zhu, and Changzhi Li, “A portable 24-ghz auditory radar for non-contact speech sensing with background noise rejection and directional discrimination,” in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, May 2016, pp. 1–4.

[9] S. Li, Y. Tian, G. Lu, Y. Zhang, H. Lv, X. Yu, H. Xue, H. Zhang, J. Wang, and X. Jing, “A 94-GHz millimeter-wave sensor for speech signal acquisition,” *Sensors*, vol. 13, no. 11, pp. 14 248–14 260, oct 2013.

[10] R. Khanna, D. Oh, and Y. Kim, “Through-wall remote human voice recognition using doppler radar with transfer learning,” *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4571–4576, jun 2019.

[11] M. Skolnik, *Radar Handbook, Third Edition*, ser. Electronics electrical engineering. McGraw-Hill Education, 2008.

[12] Y. Xiong, Z. Peng, G. Xing, W. Zhang, and G. Meng, “Accurate and robust displacement measurement for fmcw radar vibration monitoring,” *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1131–1139, Feb 2018.

[13] P. Welch, “The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” vol. 15, pp. 70–73, 1967.

[14] M. Gardill, J. Schwendner, and J. Fuchs, “An approach to over-the-air synchronization of commercial chirp-sequence automotive radar sensors,” in *2020 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT)*. IEEE, jan 2020.