RF Microphone based on Short-Range Millimeter-Wave Sensing

Li Wen, *Student Member, IEEE*, Changzhan Gu, *Senior Member, IEEE*

***Abstract*—Based on a compact custom-designed 120 GHz interferometric radar system, RF microphone as an audio recovery technique is proposed in this project. The millimeter-wave 120 GHz provides a method to precisely track a minimum of micrometer-scale displacements of vibrating sources such as working speakers as well as other objects vibrated by sound wave in real time. Experiments have been launched in office environment for functional verification of the proposed RF microphone, recovering audio signals from speakers and a hanging glass. The results show that the proposed technique performs well in sensing precise vibrations. Audio signals including single tones and musics can be clearly recovered, having a maximum coherence over 80% with the simultaneously acquired microphone signals in the frequency domain. With advantages in small size, high precision and contact-free devices, the proposed technique makes it possible for eavesdropping and surveillance of targets obstructed by barriers.**

***Index Terms*—radar, millimeter wave, audio recovery.**

I. Introduction

A microphone is the most universally used device that converts sounds to electrical signals. However, it's hard for microphones to acquire sound from afar or blocked sound sources. Therefore, the idea of using other devices and approaches to recover audio signals from surface vibrations came up, giving rise to important applications in security and surveillance. The most famous one is the visual microphone [1]. Though the approach advances in its no restriction on the observing objects, it still has obvious disadvantages, including high power consumption, long processing time, and not-real-time recovery [2]. Laser microphone [3] as another popular way to eavesdrop, can achieve real-time recovery, but has a low robustness against ambient light and needs high maintenance and line-of-sight contact [4]. In this project, an audio recovery technique based on a compact custom-designed 120 GHz interferometric radar system is proposed. The proposed technique is superior from the perspective of low power consumption, miniature size, and high accuracy.

This work was supported in part by the National Key R&D Program of China under Grant 2019YFB2204604, Shanghai Municipal Science and Technology Major Project under Grant 2021SHZDZX0102, and National Science Foundation of China under Grant 61831016. (Corresponding author: Changzhan Gu).

L. Wen and C. Gu are with the MoE Key Laboratory of High-Speed Electronic Systems and the MoE Key Laboratory of Artificial Intelligence, Shanghai Jiao Tong University, Shanghai, China, 200240 (email: changzhan@sjtu.edu.cn).

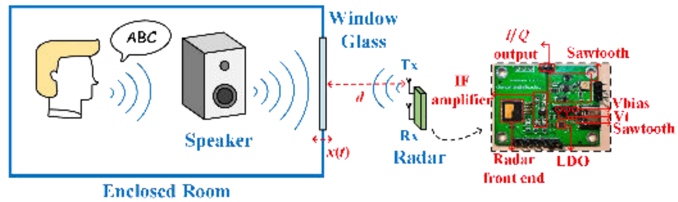

Fig. 1 Schematic diagram of the proposed RF microphone. Insert: picture of the custom designed millimeter-wave 120 GHz radar sensor system.

II. System and Theory

While the principle of traditional microphones is converting the motions of an internal diaphragm into electrical signals, the proposed RF microphone works similarly but instead detects the motions of a distant object, using the object as an external diaphragm. Fig. 1 is a schematic diagram of the proposed RF microphone in a real-life application scenario. The radar sensor continuously transmits millimeter-wave signals that are tuned to focus on a window glass of an enclosed room, where people speak or play audio files inside but all sounds are proofed. The surface vibrations of the window glass modulate the millimeter-waves in the phase and then part of the signals are backscattered back to the radar receiver.

The custom-designed 120 GHz radar sensor inserting in Fig. 1 has two working modes of continuous-wave (CW) and frequency modulated continuous wave (FMCW), the size of which is 3.24 cm by 4.27 cm. With a radar transceiver TRX_120_001 from Silicon Radar GmbH, the sensor has a frequency range from 125.9 GHz to 119.1 GHz.

As sound information is associated with the subtle surface vibrations, accurate sensing of the objects' displacements is the key in audio recovery. The wavelength of 120 GHz is only 2.5 mm, so it may easily lead to phase ambiguity as the displacements of the vibrating object could exceed half a wavelength [6]. Therefore, the modified differential and cross multiply (MDACM) algorithm [7] is applied to linearly demodulate the phase information. The relative displacement can be expressed in discrete form as

$$X[n] = \frac{\lambda}{4\pi} \sum_{k=2}^{n} I[k-1]\, Q[k] - I[k]Q[k-1] \qquad (1)$$

It is noted that long-period I and Q signals are segmented to be demodulated with a sliding window, so as to improve the demodulation accuracy and avoid the unwanted phase drift. Once the displacements of the vibrating object are obtained, two signal processing steps are employed to convert them into

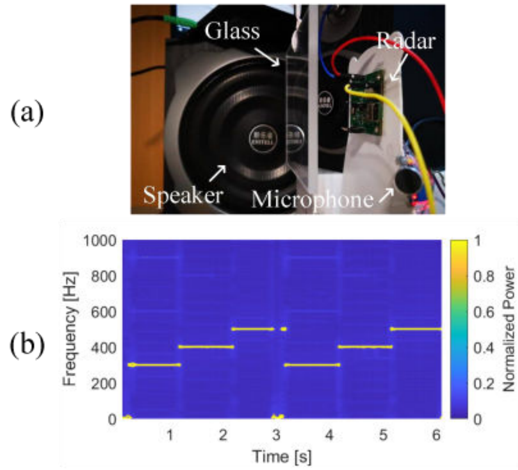

Fig. 2 (a) the experiment setup; (b) the spectrogram of the glass displacement when playing single tone signals at 300Hz, 400Hz, and 500Hz.

audio signals: signal enhancement and noise reduction. To optimize the overall perceptual quality, the obtained displacements is enhanced by normalizing the amplitude to the range of [-1,1]. Besides, spectral subtraction, one of the first techniques proposed for denoising single channel speech, is applied for a cleaner signal. Other processing methods such as machine learning can also be employed to improve the intelligibility of the signals.

III. Experiment and Result

Experiments were carried out in office environment, where a hanging glass near a speaker is set to simulate vibrating window glasses near a sound source in the real situation as shown in Fig. 2 (a). The radar sensor was placed about 5 cm away directing at the hanging glass. The radar sensor system was connected to a data acquisition (DAQ) device jointly with a microphone to synchronously obtain real-time signals.

In the first set of experiments, single tone signals were tested. Three single tone signals, which are 300Hz, 400Hz and 500Hz, were played in turn periodically by the speaker. Meanwhile, the radar sensor system detected the surface vibration of the hanging glass and converted it to voltage signals for post-processing steps. Fig. 2 (b) presents the spectrogram of the demodulated result of the glass's displacement, which matched the exact frequencies of the single tone signals.

In the second set of the experiments, a 13-second piece of music was tested, which is a track from a well-known song called "Jingle Bell" sang by Black Duck Chorus. Fig. 3 (a), (b), and (c) respectively present the spectrograms of the traditional microphone signal, the displacement of the speaker, and the displacement of the hanging glass detected by RF microphone. The lyrics shown above the figure are labeled with roman numerals, mapping with the corresponding section of the spectrograms. As shown, the displacements of the speaker and the hanging glass are basically the same, which means that the surface vibrations of the nearby object can sufficiently reflect the displacement information of the sound source. And the displacement information recovered by the RF microphone in spectrograms also has a similar pattern with the microphone signals. The coherence of the normalized microphone signal

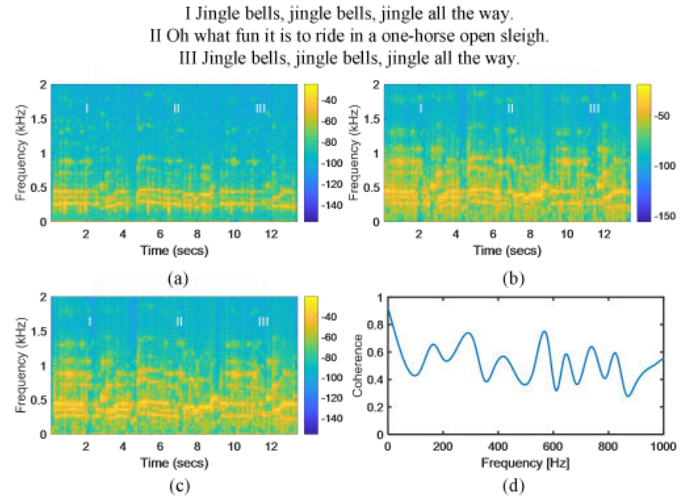

Fig. 3 The experiment results of a piece of music: (a) traditional microphone signal spectrogram; (b) the displacement spectrogram of the speaker; (c) the displacement spectrogram of the glass; (d) the coherence of the signals detected from the traditional microphone and the RF microphone.

and the glass's displacement in frequency domain is shown in Fig. 3 (d), illustrating a great similarity at low frequency with a maximum coherence over 80%. The recovered displacement not only can be visually recognized but also can be acoustically identified by automatic song identifier applications, including NetEase Music, QQ Music and SoundHound.

IV. Conclusion

A more detailed version of the final project report refers to [8]. At present, I am continuing my study for M.Sc. degree at Shanghai Jiao Tong University with my supervisor Professor Gu, and I plan to develop my work on short-range millimeter sensing further in the future. The MTT-S Scholarship has greatly inspired my confidence to take a leap into academic research, and attending IMS is also a precious experience for me to meet the experts and learn from their outstanding ideas.

References

[1] A. Davis, M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman, "The visual microphone: passive recovery of sound from video," *ACM Trans. Graph.* 33, 4, Article 79, 2014.

[2] Nassi, Ben, Yaron Pirutin, A. Shamir, Yuval Elovici and B. Zadov, " Lamphone: Real-Time Passive Sound Recovery from Light Bulb Vibrations," IACR Cryptol. ePrint Arch. 2020 (2020): 708.

[3] R. P. Muscatell, "Laser microphone," U.S. Patent 4 412 105, 1983.

[4] E. Guerrero, J. Brugués, J. Verdú and P. d. Paco, "Microwave Microphone Using a General Purpose 24-GHz FMCW Radar," in *IEEE Sensors Letters*, vol. 4, no. 6, pp. 1-4, June 2020, Art no. 3501004.

[5] Jennifer A. Rice, Changzhan Gu, Changzhi Li, Shanyue Guan, "Aradar-based sensor network for bridge displacement measurements," *Proc. Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 2012, vol. 8345.

[6] Seoktae Kim and Cam Nguyen, "A displacement measurement technique using millimeter-wave interferometry," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 6, pp. 1724-1728, 2003.

[7] W. Xu, C. Gu and J. -F. Mao, "Noncontact High-Linear Motion Sensing Based on A Modified Differentiate and Cross-Multiply Algorithm," 2020 IEEE/MTT-S International Microwave Symposium (IMS), 2020, pp. 619-622.

[8] Li Wen, Yuchen Li, Yangtao Ye, Changzhan Gu and Jun-Fa Mao, "Audio Recovery via Noncontact Vibration Detection with 120 GHz Millimeter-Wave Radar Sensing," International Conference on Microwave and Millimeter Wave Technology (ICMMT), 2021.