

## 太赫兹雷达多声音信号活动检测技术\*

王元昊,王宏强,杨琪,邓彬,罗成高  
(国防科技大学电子科学学院,湖南长沙410073)

**摘要:**针对在环境中同时存在多个声源的场景下,传统声音获取方式难以正确检测到期望声源处声音活动的问题,提出一种基于太赫兹雷达的多声音信号活动检测方法,该方法利用太赫兹雷达波长短提升对于微小振动的探测灵敏度,同时利用大宽带所带来的高距离分辨力区分不同声源,在获得声音信号后进一步通过谱减-小波联合去噪方法提升信号质量用于后续声音信号活动检测。实验结果表明,在传输路径上存在干扰声源时,相比麦克风方式而言,所提太赫兹雷达声音信号活动检测方法可以更有效地区分多个声源,获得期望声源处的声音活动。

**关键词:**太赫兹雷达;多声音活动检测;微多普勒;抗干扰

**中图分类号:**TN95 **文献标志码:**A **文章编号:**1001-2486(2022)01-028-06

## Multi-voice signal activities detection method based on terahertz radar

WANG Yuanhao, WANG Hongqiang, YANG Qi, DENG Bin, LUO Chenggao

(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)

**Abstract:** Aiming at the problem the traditional voice acquisition method is difficult to correctly detect the voice activity at the desired voice source when multiple voice sources exist simultaneously, a multi-voice activities detection method based on terahertz radar was proposed. The short wave length of terahertz radar was utilized to improve the detection sensitivity for micro vibration. At the same time, the high range resolution brought by large broadband was used to distinguish different voice sources. After obtaining the voice signal, the signal quality was further improved for signal activities detection by spectral subtraction wavelet joint denoising method. The experimental results show that, when there are interfering voice source on the transmission path, the proposed method can more effectively distinguish multiple voice sources and obtain the voice activity at the desired voice source than the microphone-based method.

**Keywords:** terahertz radar; multi-voice activities detection; micro-Doppler; anti-interference

太赫兹指的是频率在0.1~10 THz之间的电磁波,太赫兹雷达就是指工作在太赫兹频段下的雷达。相比于传统微波频段雷达,太赫兹雷达在精度方面具有天然的优势,使其在拓展传统雷达应用领域上具有巨大的潜力。声音活动检测(Voice Activity Detection, VAD)指从包含声音的信号中确定出声音的起点及终点,即确定待处理信号是否为有效声音信号,是声音信号预处理的重要一环。对于后续的声音增强,声音识别等处理具有重要意义。目前的VAD算法多采取基于信号特征的形式进行声音或非声音的判别,常用的特征有能量、频域、倒谱、谐波等。但这些特征都依赖于所获得信号中的噪声分量具有和声音分量显著的特征差异,基于深度学习的VAD方法也

基本建立在特征值的提取上<sup>[1-2]</sup>,显著的特征差异对于识别算法起着至关重要的作用,获取声音信号的方式对于特征值的获取也有着至关重要的作用。

目前,麦克风方式是发展最为成熟<sup>[3]</sup>、应用最为广泛的一种声音信号获取方式,但是利用麦克风存在一些弊端,导致其应用受限:一是抗干扰能力差,即面对所期望得到的声音湮没在噪声中或传输路径中存在干扰声源时该方法难以取得良好的性能。二是作用距离近,因为声音信号在空气中的衰减较大,所以一般都会放置在声源米量级内的位置来获得最佳的效果。但若在麦克风接收路径上存在其他声源的声音干扰,则即便是信道噪声不强,在进行VAD时也会存在较大的误

\* 收稿日期:2021-01-25

基金项目:国家自然科学基金资助项目(61871386,61971427)

作者简介:王元昊(1997-),男,河南新乡人,博士研究生,E-mail:playsunsky@163.com;

杨琪(通信作者),男,副教授,博士,E-mail:yangqi08@nudt.edu.cn

差。因此,探索一种新的 VAD 手段具有重要意义。探测声波本质上是探测物体的振动信息,而雷达天然地具有测距能力。王健琪团队在 2016 年提出基于 94 GHz 雷达的生物声音获取系统<sup>[3-4]</sup>,利用毫米波雷达对人声带振动信号进行远程检测从而得到声音信号。Khanna 等在 2019 年提出基于多普勒雷达的声音识别系统<sup>[5]</sup>,该系统放置在离人咽喉很近的地方。相比于麦克风,雷达获取声音信号的方式可以认为是远程直接探测声源表面的振动,可以有效地避免传输路径上其余声音干扰带来的影响。但是上述对于雷达声音获取技术的研究仍然存在不具备目标选择能力以及多目标同时获取能力不足的问题。

为此,本文提出一种基于载波频率为 120 GHz 的宽带太赫兹雷达的 VAD 技术,综合利用太赫兹雷达频段高和带宽大的特性,实现多声音信号活动检测,同时通过谱减-小波联合去噪方法提升雷达获取声音信号的质量。

## 1 太赫兹雷达声音活动探测原理

### 1.1 声音信号传播模型

声音信号本质上是由物体振动产生的机械波。假设一个无向声源,该声源的振动方程为  $f(x, t)|_{x=0}$ 。根据傅里叶变换原理,声音信号可以分解为不同频率、幅度的正弦信号的线性叠加,因此  $f(x, t)|_{x=0}$  可以重写为:

$$f(x, t)|_{x=0} = \int A_{\omega} \cos(\omega t + \varphi_{\omega}) d\omega \quad (1)$$

式中,  $A_{\omega}$  是频率分量为  $\omega$  的谐波所对应的振幅,  $\varphi_{\omega}$  是其相应的相位值。

考虑  $M$  个主要频率分量以及因路径传播所带来的延迟,式(1)可以重写为:

$$F(x, t) = \sum_{i=1}^M A_{\omega_i} \cos\left[\omega_i \left(t - \frac{x}{\lambda}\right) + \varphi_{\omega_i}\right] \quad (2)$$

式中,  $A_{\omega_i}$  是频率分量为  $\omega_i$  的谐波所对应的振幅,  $\varphi_{\omega_i}$  是对应的相位,  $\lambda$  表示波长。

当声波传递到硬质介质表面时,会引发该介质的振动,考虑振动方程:

$$\frac{\partial^2 r}{\partial t^2} + 2\delta \frac{\partial r}{\partial t} + \omega_0^2 r = \frac{1}{m} F(x, t) \quad (3)$$

式中,  $\delta$  为阻尼系数,  $\omega_0$  表示物体本振频率,  $m$  表示该物体质量。

假设  $M=1$ , 求解式(3)可得:

$$r = e^{-\delta t} \left[ r_0 \cos(\omega_1 t) + \frac{\delta r_0}{\omega_1} \sin(\omega_1 t) \right] +$$

$$e^{-\delta t} A_{\omega_1} \frac{\omega_0^2 \omega_1 F}{(\omega^2 - \omega_1^2)^2 + 4\delta^2 \omega_1^2} \cdot \left[ 2\delta \cos(\omega_1 t) + \frac{2\delta^2 - \omega^2 + \omega_1^2}{\omega_1} \sin(\omega_1 t) \right] + A_{\omega_1} \frac{\omega_0^2 \omega_1 F}{(\omega^2 - \omega_1^2)^2 + 4\delta^2 \omega_1^2} \cdot \left[ (\omega^2 - \omega_1^2) \sin(\omega_1 t) - 2\delta \omega_1 \cos(\omega_1 t) \right] \quad (4)$$

式中,  $r_0$  为该点的初始位移,  $\omega = \sqrt{\omega_0^2 - \delta^2}$ 。

分析式(4)可得,含有  $e^{-\delta t}$  的前两个振动项为自由振动,会随着时间迅速衰减,而只有最后一项为受迫振动,该受迫振动项将以与声源频率相同的频率振动且振幅与声源振幅  $A$  呈线性关系。

### 1.2 太赫兹雷达微动测量原理

设太赫兹雷达发射调频连续波,记发射信号为:

$$s(t) = \exp\left[j2\pi\left(f_0 t \pm \frac{\gamma}{2} t^2\right)\right] \quad (5)$$

式中,  $f_0$  是中心频率,  $\gamma$  是调频率。

假设目标有  $N$  个散射中心,可以得到回波信号表达式为:

$$r(t) = \sum_{i=1}^N \sigma_i \exp\left\{j2\pi\left[f_0(\hat{t} - \tau_i) \pm \frac{\gamma}{2}(\hat{t} - \tau_i)^2\right] + \phi_i\right\} \quad (6)$$

式中,  $\sigma_i$  表示不同散射点的回波散射强度,  $\tau_i$  是不同散射点的延迟,  $\phi_i$  表示剩余相位项,  $\hat{t}$  表示快时间且  $\hat{t} = t \bmod T$  ( $T$  为扫频周期)。

设参考信号为:

$$s_{\text{ref}}(t) = \exp\left\{j2\pi\left[f_0\left(t - \frac{2R_{\text{ref}}}{c}\right) + \frac{\gamma}{2}\left(\hat{t} - \frac{2R_{\text{ref}}}{c}\right)^2\right]\right\} \quad (7)$$

式中,  $c$  为光速,  $R_{\text{ref}}$  为参考距离。

将回波信号  $r(t)$  和参考信号  $s_{\text{ref}}(t)$  相乘,得到去斜后信号:

$$r'(t) = \exp\left[-j\left(\frac{4\pi f_0}{c} R_{\Delta} + \frac{4\pi\gamma}{c^2} R_{\Delta}^2 + \frac{4\pi f_i}{c} R_{\Delta}\right)\right] \quad (8)$$

式中,  $R_{\Delta} = R_i - R_{\text{ref}}$ ,  $R_i = \frac{\tau_i c}{2}$ 。

进一步利用相位测距<sup>[6-7]</sup>原理可得到目标位移如下:

$$\Delta\phi = \text{angle}[r'(t + \Delta t) - r'(t)] = 2\pi f_0 \Delta\tau = \frac{4\pi f_0}{c} \Delta R \quad (9)$$

$$\Delta R = -\frac{c}{4\pi f_0} \Delta\phi = -\frac{\lambda}{4\pi} \Delta\phi \quad (10)$$

在得到位移变化序列后进一步通过联合谱减

算法<sup>[8-9]</sup>以及小波去噪<sup>[10]</sup>提升信号质量。

谱减法是从含噪信号功率谱中减去噪声功率谱,要求噪声与纯净声音信号不相关。纯净声音信号与噪声的相关性决定了去噪后声音信号的质量。然而,由于有色噪声的分布特点,去除噪声的效果并不理想。但是有色噪声的功率很低,可以采用小波去噪的方法。因为小波去噪的基本思想是对信号进行小波变换,通过选择合适的阈值来保留较大的小波系数,从而保留了信号的主要分量。

记通过相位测距处理得到的位移序列  $y[n] = \{R(t), R(t+T), \dots, R(t+mT)\}$ , 其可以认为是含噪信号,假设纯净声音信号和噪声是不相关的。

含噪信号可以表示为:

$$y[n] = \begin{cases} d_{\text{wgn}}[n], & \text{不存在声音信号} \\ x[n] + d_{\text{cgn}}[n] + d_{\text{wgn}}[n], & \text{存在声音信号} \end{cases} \quad (11)$$

其中,  $x(n)$  表示干净信号,  $d_{\text{cgn}}(n)$  表示有色噪声,  $d_{\text{wgn}}(n)$  表示白噪声。

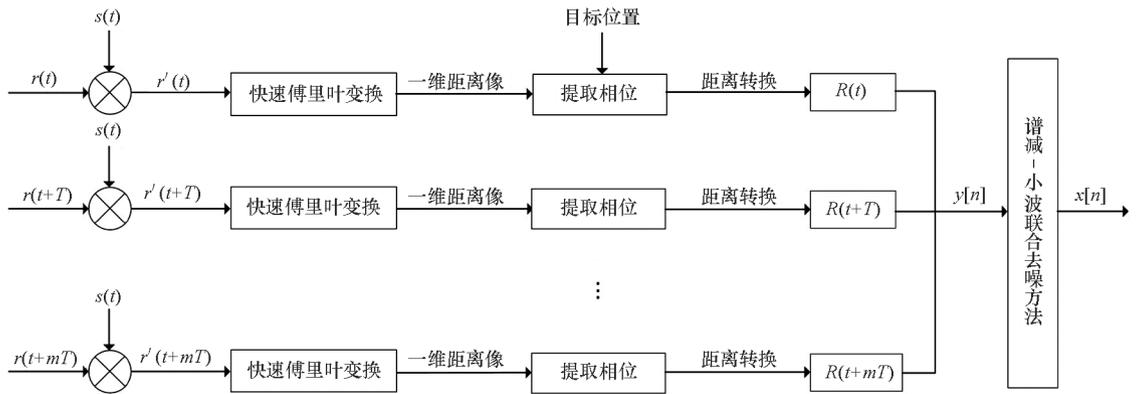


图 1 太赫兹雷达声音活动检测信号处理框图

Fig. 1 Signal processing flow of VAD system for terahertz radar

## 2 太赫兹雷达声音探测系统

本文所设计太赫兹雷达声音活动检测系统主要由原始数据获取前端和数据精细化处理后端组成。其结构框图如图 2 所示。原始数据获取前端主要是一部太赫兹雷达,其发射波形为调频连续波,中心频率为 120 GHz,带宽为 0.1~4 GHz(可调节),扫频周期为 0.1~1 ms(可调节)。数据精细化处理后端主要由一台上位机组成,其作用是控制雷达参数、接收雷达数据并处理。

因为雷达的距离分辨率为  $R_{\text{res}} = c/(2B)$ , 带宽  $B$  选择为 500 MHz 可以区分目标声源与干扰声源。扫频周期选择为 0.5 ms,扫频周期的倒数对应于实际对声源信号的采样率,0.5 ms 对应采

将式(11)转换到频域,可以表示为:

$$Y[k] = \begin{cases} D_{\text{wgn}}[k], & \text{不存在声音信号} \\ X[k] + D_{\text{cgn}}[k] + D_{\text{wgn}}[k], & \text{存在声音信号} \end{cases} \quad (12)$$

其中,  $Y[k]$ 、 $X[k]$ 、 $D_{\text{cgn}}[k]$  和  $D_{\text{wgn}}[k]$  分别是  $y[n]$ 、 $x[n]$ 、 $d_{\text{cgn}}(n)$  和  $d_{\text{wgn}}(n)$  的傅里叶变换。

进一步转换成功率谱可得:

$$|Y[k]|^2 = \begin{cases} |D_{\text{wgn}}[k]|^2, & \text{不存在声音信号} \\ |X[k]|^2 + |D_{\text{cgn}}[k]|^2 + |D_{\text{wgn}}[k]|^2, & \text{存在声音信号} \end{cases} \quad (13)$$

通过选择不含声音信号的一部分来估计纯净信号:

$$|\hat{X}[k]|^2 = |Y[k]|^2 - |\hat{D}_{\text{wgn}}[k]|^2 \quad (14)$$

式中,  $\hat{D}_{\text{wgn}}[k]$  是估计出的白噪声功率谱。

最后,通过小波降噪的方法进一步去除  $d_{\text{cgn}}[n]$ ,从而得到干净声音信号  $x[n]$ 。信号处理框图如图 1 所示。

样率为 2 000 Hz,可以满足对于检测声音活动的需求。

## 3 系统设计与实验验证

设计实验场景如图 3 所示,实验环境为微波暗室和户外,选取两个蓝牙音箱分别作为目标声源和干扰声源,目标声源播放声音“电子科学学院”,声音信号获取系统选择麦克风和太赫兹雷达,麦克风系统主要用来与太赫兹雷达对比所获得声音信号的 VAD 性能。

实验室环境下太赫兹雷达测得的声音信号如图 4 所示。对于目标声源,雷达通过相位测距方式得到的原始信号、经过谱减法降噪后的声音信号、谱减法与小波去噪结合降噪后的声音信号的

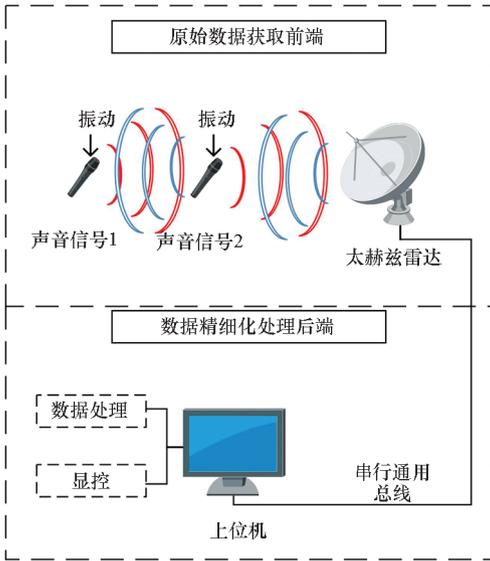
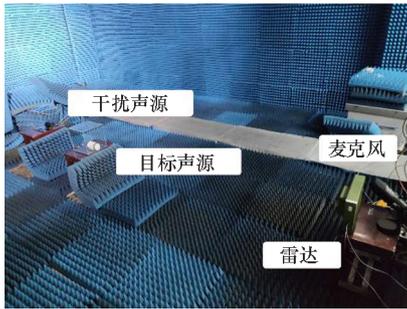


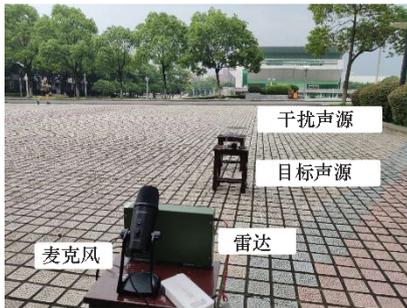
图 2 太赫兹雷达声音活动检测系统框图

Fig. 2 Block diagram of VAD system for terahertz radar



(a) 实验室环境

(a) Laboratory environment



(b) 户外环境

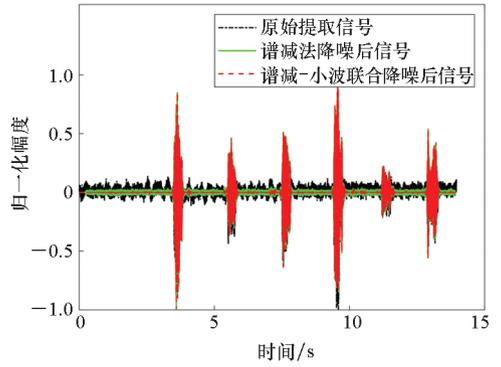
(b) Outdoor environment

图 3 实验场景

Fig. 3 Experimental environment

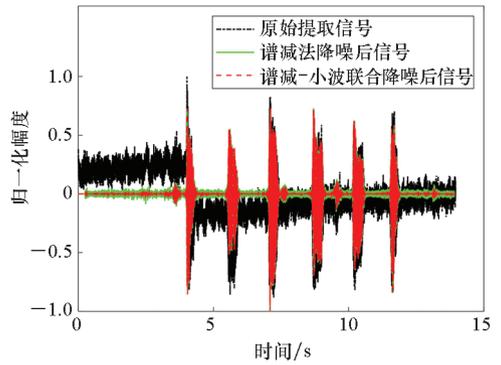
信噪比分别为 7.53 dB、20.69 dB 和 36.55 dB。对于干扰声源, 该数值分别为 9.69 dB、20.10 dB 和 29.38 dB。从图 4 和计算的信噪比可以看出, 谱减-小波联合降噪算法具有明显的信噪比改善。

同样给出户外测试的结果, 如图 5 所示。从图 5 中可以看出, 相比于实验室环境, 户外环境下原始提取信号的幅度抖动十分剧烈, 通过所提谱减-小波联合降噪方法可以有效抑制噪声, 保留声音信号。



(a) 目标声源

(a) Desired speech signal



(b) 干扰声源

(b) Interference speech signal

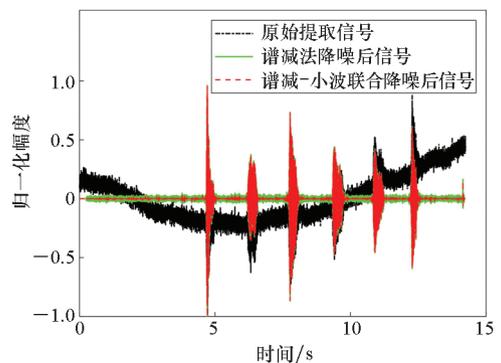
图 4 实验室环境下的太赫兹雷达探测声音信号结果

Fig. 4 Results of acoustic signal acquired by terahertz radar in laboratory environment

麦克风在实验室环境中采集的声音信号的结果如图 6 所示。由于麦克风不能区分不同位置处的声源, 采集到的声音信号存在严重混叠现象。

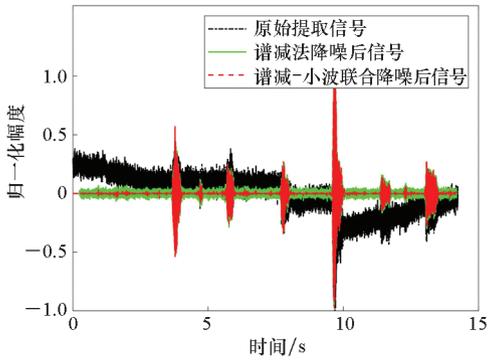
从图 6 中可以看出, 麦克风所采集声音信号无法区分声音信号来源是目标物体还是干扰物体; 当在传输信道中施加干扰声源后, 麦克风采集声音信号难以正确判断目标物体是否进行声音活动。

图 7 展示了基于太赫兹雷达所采集声音信号



(a) 目标声源

(a) Desired speech signal



(b) 干扰声源

(b) Interference speech signal

图 5 户外环境下的太赫兹雷达探测声音信号结果  
Fig. 5 Results of acoustic signal acquired by terahertz radar in outdoor environment

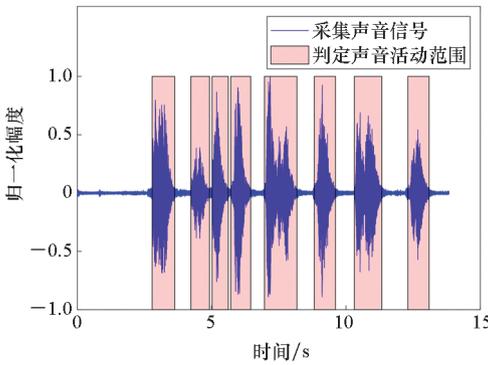


图 6 麦克风采集声音信号 VAD 结果  
Fig. 6 VAD result of microphone

的 VAD 结果。因为雷达天然地具有距离分辨的能力,所以首先通过距离分辨选择目标所在的距离单元数据进行进一步分析,可以有效地避免采集来自干扰声源的信号。通过对比图 6 可以明显看出,基于太赫兹雷达采集的声音信号的 VAD 效果明显优于麦克风方式。

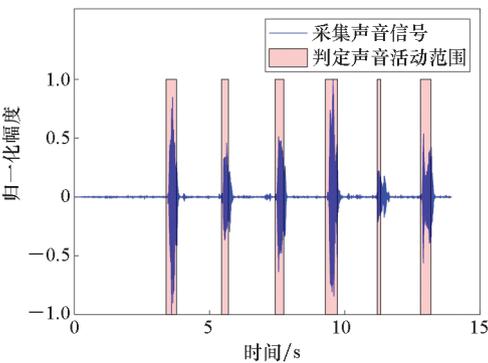


图 7 雷达采集声音信号 VAD 结果  
Fig. 7 VAD result of terahertz radar

进一步,通过声源与雷达的距离来对该系统的应用场景进行评估,以信噪比作为评判标准,捕捉最初无声音信号播放时的雷达数据作为噪声样

本,得到结果如图 8 所示。

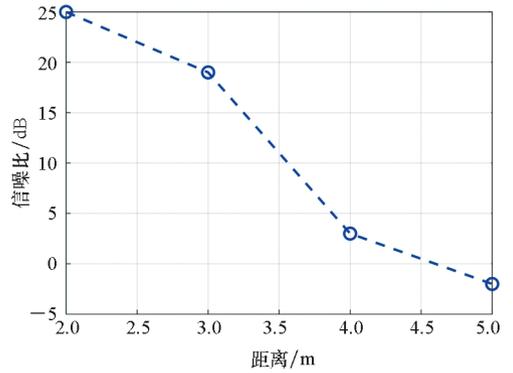


图 8 探测声音信号信噪比与作用距离的关系  
Fig. 8 Relationship between signal to noise ratio and operating distance of sound signal

从图 8 中可以看出,该系统的探测性能随距离下降十分明显,但该性能的下落可以通过提升雷达功率的方式来得以改善。

### 4 结论

本文提出了一种太赫兹雷达多声音检测方法。太赫兹雷达凭借其频率高、波长短的特点,在相位测距精度上有着显著的优势。本文采用太赫兹雷达进行声音信号获取并进行声音活动检测,同时在相同环境下也设置了麦克风进行声音信号获取来进行对比。实验结果表明,太赫兹雷达声音检测方法在面临传输路径上存在声源干扰时,对目标声源的声音活动仍然有良好的判断性能,在一定程度上可以弥补基于麦克风的语音信号采集系统的不足,该方法为在复杂环境下语音信号的探测提供了一种新的思路。

### 参考文献 (References)

[1] DOV D, TALMON R, COHEN I. Audio-visual voice activity detection using diffusion maps[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015, 23(4): 732-745.

[2] KRISHNAKUMAR H, WILLIAMSON D S. A comparison of boosted deep neural networks for voice activity detection[C]//Proceedings of IEEE Global Conference on Signal and Information Processing, 2019.

[3] 陈扶明, 李盛, 安强, 等. 生物雷达语音信号探测技术研究进展[J]. 雷达学报, 2016, 5(5): 477-486.

CHEN F M, LI S, AN Q, et al. Advancements in bio-radar speech signal detection technology[J]. Journal of Radars, 2016, 5(5): 477-486. (in Chinese)

[4] CHEN F M, LI S, ZHANG Y, et al. Detection of the

- vibration signal from human vocal folds using a 94-GHz millimeter-wave radar[J]. *Sensors*, 2017, 17(3): 543.
- [5] KHANNA R, OH D G, KIM Y W. Through-wall remote human voice recognition using Doppler radar with transfer learning[J]. *IEEE Sensors Journal*, 2019, 19(12): 4571 – 4576.
- [6] ZHU D K, LIU Y X, HUO K, et al. A novel high-precision phase-derived-range method for direct sampling LFM radar[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(2): 1131 – 1141.
- [7] FAN H Y, REN L X, MAO E K. A micro-motion measurement method based on wideband radar phase derived ranging [ C ]//Proceedings of IET International Radar Conference, 2013.
- [8] ISLAM M T, SHAHNAZ C, FATTAH S A. Speech enhancement based on a modified spectral subtraction method[ C ]//Proceedings of IEEE 57th International Midwest Symposium on Circuits and Systems, 2014: 1085 – 1088.
- [9] 胡金艳, 李盛, 陈扶明, 等. 基于多带谱减法的生物雷达语音增强方法研究 [ J ]. 科学技术与工程, 2017, 17(16): 76 – 81.
- HU J Y, LI S, CHEN F M, et al. A new kind of nonlinear spectral subtraction method for radar speech enhancement [ J ]. *Science Technology and Engineering*, 2017, 17(16): 76 – 81. (in Chinese)
- [10] CHEN F M, LI C T, AN Q, et al. Noise suppression in 94 GHz radar-detected speech based on perceptual wavelet packet[ J ]. *Entropy*, 2016, 18(7): 265.