

Improved Throat Vibration Sensing with a Flexible 160-GHz Radar through Harmonic Generation

Martin Geiger, Denis Schlotthauer, and Christian Waldschmidt

Improved Throat Vibration Sensing with a Flexible 160-GHz Radar through Harmonic Generation

Martin Geiger, Denis Schlotthauer, and Christian Waldschmidt
Institute of Microwave Engineering, Ulm University, 89081 Ulm, Germany
Email: martin-2.geiger@uni-ulm.de

Abstract—Speech can be recorded contact-free using the Doppler effect with a radar microphone and comprehensible results can be achieved even in a noisy environment. For that purpose, the transmit frequency must be higher than 100 GHz to measure enough harmonics of the skin vibration. This paper presents a 160 GHz radar MMIC with a transition to a flexible dielectric waveguide and a plugged antenna for a radar microphone application. With the flexible antenna front-end made of high density polyethylene a comfortable and easy attachment similar to an ordinary microphone is possible. The measured speech signals are comprehensible with the fundamental frequency and at least three harmonics. The background noise does not affect the speech signal as proven in measurements.

Index Terms—radar microphone, millimeter wave, Doppler radar, flexible antenna.

I. INTRODUCTION

The fundamental part of communication is speech. It is generated by the vibration of the vocal chords and resonances in the vocal tract [1]. The vocal chords are excited by air pressure and resonate with the fundamental speech frequency and its harmonics, which decrease with 12 dB per octave. These vibrations propagate to the skin surface and can be sensed. The formants are then resonances in the vocal tract with tongue, lips, and jaw.

To detect these vibrations an audio microphone is usually used. However, in urban areas or in factory environments with high background noise complex signal processing algorithms are needed to extract the spoken words. Non-acoustic speech sensors like general electromagnetic motion sensors, physiological sensors, or bone-conduction microphones are immune against background noise [2] but are attached to the skin of the speaker and can restrict the flexibility or cause skin irritations.

A possible contact-free measurement method would be a radar system. By analyzing the Doppler frequency a radar can detect motions or vibrations. Such devices are known from contactless heart beat measurements or other vital sign monitoring [3]. Furthermore, different radar microphones were proposed with transmit frequencies of 925 MHz [4] and 24 GHz [5]. However, these systems are bulky and are more difficult to install than existing sensors. Furthermore, apart of the fundamental frequency only one harmonic can be measured. For a comprehensible speech detection these frequency components are insufficient and additional algorithms are needed to obtain clear voice signals.

An antenna that is concentrated to a PCB with the radar sensor using a flexible waveguide can easily be clipped to a

shirt like standard microphones and can therefore satisfy many diverse applications.

To increase the quality of the recorded speech, the number of measured harmonics must be increased compared to the existing radar microphones. As the vibration amplitude is strongly varying for the different frequency components, a possible solution is to choose the transmit frequency such that the harmonics are measured as modulations of the actual measured frequency.

In this paper, a radar microphone with a flexible front-end is presented. The operation frequency of 160 GHz makes it possible to measure several harmonics. At first the vibration detection theory is explained. Afterwards, the used system is presented and the measurement results for a spoken sentence are shown.

II. SPEECH DETECTION THEORY

The basic principle behind speech detection with a radar is the Doppler effect. A continuous wave radar (CW radar) transmits a continuous signal with frequency f_t . Reflected signals are received and down-converted with the phase-shifted transmit signal. The down-converted signal contains information about the distance R between the throat and the radar, the vibration frequency f_v , and is given by

$$u_{\text{IF}}(t) = \hat{u}_{\text{IF}} \cos\left(\frac{4\pi x(t)}{\lambda} + \phi_r\right), \quad (1)$$

where $x(t)$ is the time dependent vibration, $\lambda = c_0/f_t$ the wavelength, $\phi_r = \frac{4\pi f_t R}{c_0}$ the range dependent phase, and \hat{u}_{IF} the amplitude. For a single-frequency vibration with amplitude A_v , the vibration $x(t) = A_v \sin(2\pi f_v t)$ and (1) becomes

$$u_{\text{IF}}(t) = \hat{u}_{\text{IF}} \cos\left(4\pi \frac{A_v}{\lambda} \sin(2\pi f_v t) + \phi_r\right).$$

This equation can be decomposed into its frequency components [4]:

$$u_{\text{IF}}(t) = \hat{u}_{\text{IF}} \left[J_0\left(4\pi \frac{A_v}{\lambda}\right) \cos \phi_r - 2 \sum_{k=1}^{\infty} J_{2k-1}\left(4\pi \frac{A_v}{\lambda}\right) \sin \phi_r \sin[(2k-1)2\pi f_v t] + 2 \sum_{k=1}^{\infty} J_{2k}\left(4\pi \frac{A_v}{\lambda}\right) \cos \phi_r \cos(2k \cdot 2\pi f_v t) \right], \quad (2)$$

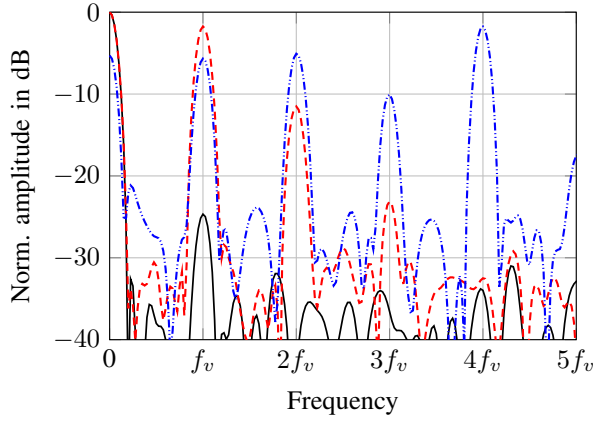


Fig. 1. IF signal spectrum of a vibration with frequency f_v for $A_v/\lambda = 0.01$ (—), $A_v/\lambda = 0.1$ (---), and $A_v/\lambda = 1$ (-.-.-).

where J_n is the n^{th} -order Bessel function of the first kind. Equation (2) contains the vibration frequency as well as the harmonics with different amplitude weightings.

One decisive amplitude term is $J_n(4\pi \frac{A_v}{\lambda})$. With increasing wavelength, the number of measured harmonics decreases for a vibration with constant amplitude, which is shown in Fig. 1. The second amplitude term is $\cos \phi_r$ for the odd and $\sin \phi_r$ for the even components. They become zero for $\phi_r = \pi/2$ respectively $\phi_r = \pi$.

For the correct frequency detection of a single frequency vibration, the harmonics should be suppressed and only the fundamental frequency should be detected. This can be achieved with $A_v/\lambda \leq 0.01$ [4].

Since the voice is a multitone vibration with a fundamental frequency around 100 Hz to 200 Hz and several harmonics up to more than 5 kHz, the harmonic suppression is a disadvantage. The detection of harmonics for a throat microphone is necessary, as only the fundamental frequency with an amplitude $A_v < 0.5$ mm and the first few harmonics with even smaller amplitudes are stimulated in the glottis. The formants which are excited in the vocal tract are not measurably with the throat microphone. Consequently, a low harmonic suppression is desirable for the detection of speech with a radar microphone. Thus, A_v/λ should be between 0.1 and 1. The radar frequency is then higher than 100 GHz.

The problem of the extinction of harmonics due to ϕ_r does not exist for these frequencies. For a displacement of $460 \mu\text{m}$ the phase change is $\Delta\phi_r = \pi$ at 160 GHz. As the throat is an extensive target, different values for ϕ_r are measured.

III. SYSTEM DESCRIPTION

In addition to a transmit frequency higher than 100 GHz the used radar should have a flexible front-end for a simple and comfortable attachment near the throat.

The used radar system works at 160 GHz with a flexible lens antenna. The system concept is shown in Fig. 2 and is explained in more detail in [6]. The bistatic radar MMIC has low phase noise due to its frequency offset synthesizer and can

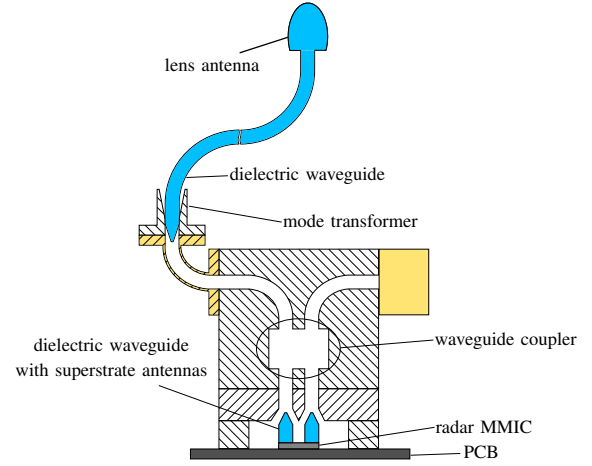


Fig. 2. System concept of the 160 GHz radar MMIC with the transition to the flexible dielectric antenna [6].

generate CW signals as well as FMCW ramps over a bandwidth of 16 GHz. The on-chip antenna radiates the signal with a $\lambda/2$ patch resonator on a quartz glass carrier into a dielectric waveguide made of acrylic glass. This dielectric waveguide is tapered and radiates into a widened metallic waveguide. A wideband hybrid coupler in rectangular waveguide technology provides the transition between the bistatic sensor and the monostatic antenna architecture. A modes transformer is needed for the transition from the metallic waveguide to the dielectric waveguide, the flexible part of the system. Since the fundamental mode of the both waveguides, the TE_{10} mode and the HE_{11} mode are very similar, the transition is a widened metallic waveguide with a tapered dielectric waveguide in it. This dielectric waveguide is made of high density polyethylene (HDPE) ($\epsilon_r = 2.25$, $\tan \delta = 3.1 \cdot 10^{-4}$) and can be bended up to a radius of 1.5 cm with negligible radiation losses. Due to the field distribution in the dielectric waveguide, the straight waveguide with a cross-section of $648 \mu\text{m} \times 1295 \mu\text{m}$ has low losses of 4.5 dB/m. A dielectric lens antenna with 28 dBi gain and a diameter of 1.5 cm is plugged on the dielectric waveguide.

This low-loss, high gain, and flexible front-end is particularly well suited for a throat microphone. The PCB with the metallic waveguide transitions can be stowed in a pocket, and the flexible dielectric waveguide with lens antenna can be positioned as required to measure the throat vibrations. To prevent additional losses due to contact with other materials the flexible waveguide is shielded with Rohacell ($\epsilon_r \gtrsim 1$).

IV. MEASUREMENT RESULTS

For the measurement with the flexible radar microphone, the antenna was positioned 5 cm apart from the throat as shown in Fig. 3. The test person, a young man, recorded the sentence ‘‘I’m Brian. Hi.’’ with both a standard audio microphone and the radar microphone. The radar signal was processed with a short-time Fourier transform, and a spectral subtraction scheme [7] was applied to minimize the noise.



Fig. 3. Flexible positioning of the dielectric lens antenna at the throat for the vibration measurement.

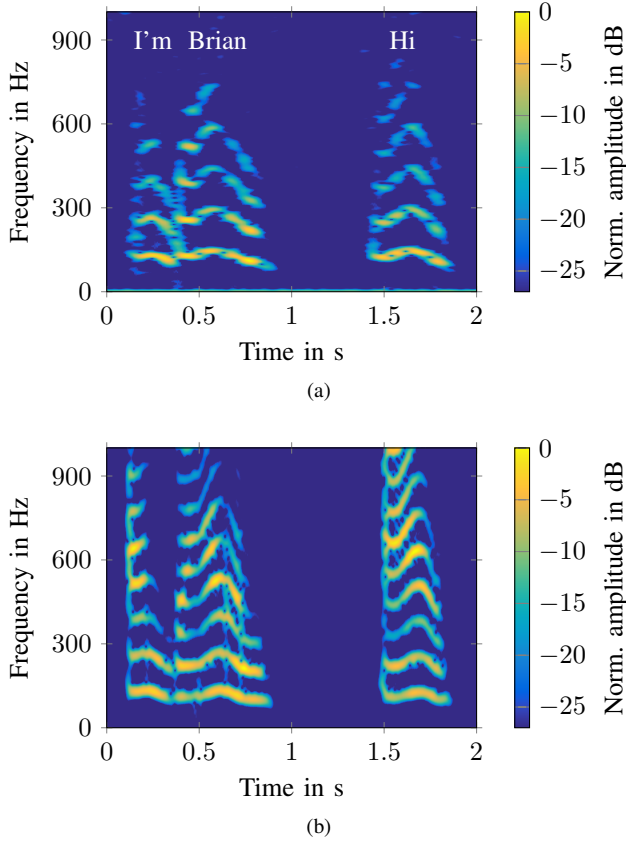


Fig. 4. Spectrogram of radar-detected (a) and audio microphone signal (b).

In Fig. 4a the spectrogram of the radar-detected signal is shown. The spectrogram consists of the three words “I’m” (0.2 s – 0.4 s), “Brian” (0.4 s – 0.9 s), and “Hi” (1.4 s – 1.8 s). The fundamental frequency is between 100 Hz and 160 Hz and depends on the spoken vowels. For every fundamental frequency at least three harmonics are measured. They are necessary for a comprehensible speech detection. Even parts of the 4th and the 5th harmonics are measured.

Compared to the audio spectrogram (Fig. 4b) less harmonics are measured since the skin only vibrates with the fundamental frequency. The higher harmonics have significant lower amplitudes or are only formed by the vocal tract, the

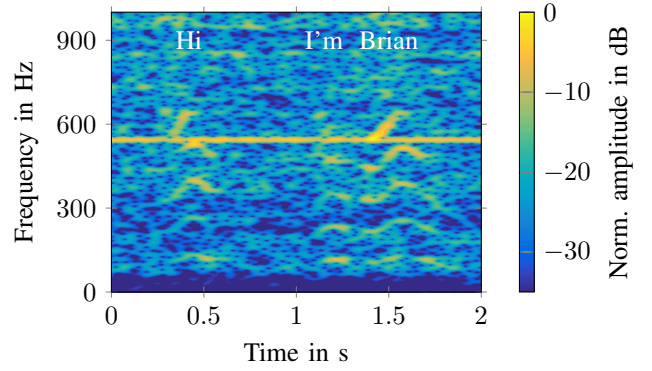


Fig. 5. Spectrogram of the audio microphone signal with a vacuum cleaner.

TABLE I
COMPARISON OF CW RADAR MICROPHONES DETECTING SPEECH.

Reference	This work	[4]	[5]
Antenna	flexible lens	beverage antenna	4×4 patch array
Frequency	160 GHz	0.925 GHz	24 GHz
A_v/λ	0.1067	0.0006	0.016
Harmonics	3	1	0

tongue, or the lips. Additionally harmonics improve the speech comprehension, however, the radar-detected speech signal is also comprehensible with three harmonics. In an acoustical comparison the audio microphone sounds clearer and cleaner and especially the consonants are more understandable.

In noisy environments the signal of the audio microphone measurement is heavily disturbed since it is sensitive to air vibrations. A measured spectrogram with a vacuum cleaner in operation next to the speaker is shown in Fig. 5. The spoken signal is almost covered by the noise spectrum of the vacuum cleaner and is almost inaudible, whereas the radar microphone is not influenced by the noise as it only measures the throat vibrations.

A comparison to other radar microphone systems is given in Table I. With the higher frequency and consequently a larger A_v/λ at least three harmonics could be detected. The flexible antenna configuration is also advantageous compared to the other antenna systems.

V. CONCLUSION

In this paper, a radar system with a flexible antenna for a radar microphone application is presented. The flexible front-end offers the possibility to place the microphone comfortably at the throat.

The radar system consists of a 160 GHz radar MMIC with a transition to a flexible dielectric waveguide with a high gain lens antenna. At the operating frequency the fundamental frequency of speech as well as at least three harmonics could be detected since the signal quality depends on the ratio A_v/λ . While standard audio microphones are sensitive to background noise, the results of the radar microphone are not affected by a noisy environment.

ACKNOWLEDGMENT

This work was supported by the Ministry for Science, Research and Arts Baden-Württemberg within the project ZAFH MikroSens.

REFERENCES

- [1] M. Redford, *The Handbook of Speech Production*, ser. Blackwell Handbooks in Linguistics. Wiley, 2015.
- [2] T. F. Quatieri, K. Brady, D. Messing, J. P. Campbell, W. M. Campbell, M. S. Brandstein, C. J. Weinstein, J. D. Tardelli, and P. D. Gatewood, "Exploiting nonacoustic sensors for speech encoding," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 2, pp. 533–544, Mar. 2006.
- [3] M. Leib, W. Menzel, B. Schleicher, and H. Schumacher, "Vital Signs Monitoring with a UWB Radar Based on a Correlation Receiver," in *Proceedings of the Fourth European Conference on Antennas and Propagation*, Apr. 2010, pp. 1–5.
- [4] C. S. Lin, S. F. Chang, C. C. Chang, and C. C. Lin, "Microwave Human Vocal Vibration Signal Detection Based on Doppler Radar Technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 8, pp. 2299–2306, Aug. 2010.
- [5] H. Zhao, Z. Peng, H. Hong, X. Zhu, and C. Li, "A Portable 24-GHz Auditory Radar for Non-contact Speech Sensing with Background Noise Rejection and Directional Discrimination," in *IEEE MTT-S International Microwave Symposium (IMS)*, May 2016, pp. 1–4.
- [6] M. Geiger, M. Hitzler, S. Saulig, J. Iberle, P. Hügler, and C. Waldschmidt, "A 160-GHz Radar With Flexible Antenna Used as a Sniffer Probe," *IEEE Sensors Journal*, vol. 17, no. 16, pp. 5104–5111, Aug. 2017.
- [7] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," in *79. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, Apr. 1979, pp. 208–211.