Microwave Human Vocal Vibration Signal Detection Based on Doppler Radar Technology

Chien-San Lin, Sheng-Fuh Chang, *Senior Member, IEEE*, Chia-Chan Chang, *Member, IEEE*, and Chun-Chi Lin, *Student Member, IEEE*

***Abstract*—A speech radar system is presented for extracting speech information from the vocal vibration signal of a human subject. Due to the tiny glottis motion of several millimeters, a coherent homodyne demodulator with high sensitivity is developed to detect reflected radio signal, phase modulated by the vibrating vocal cords. The signal detection quality and system circuit design are described. Measurements of vowels and words, both with the speech radar system and the conventional acoustic microphone system, were conducted and compared. The essential speech information can be reliably obtained from the proposed speech radar, making it more appealing for speech applications in high background acoustic noise environment.**

***Index Terms*—Coherent demodulation, Doppler radar, glottis, homodyne demodulation, vocal cords.**

I. Introduction

IT HAS been acknowledged that speech is a most effective method of human communication. Since the 1960s, scientists and engineers have made significant progress in speech signal processing, such as speech coding and synthesizing, speech recognition, speaker verification and identification, etc. Most techniques are based on the data of spoken voices. However, the recorded acoustic signals easily compose the surrounding background noise, which will considerably degrade the signal quality.

In physiology, instruments including the electroglottograph [1] and high-speed video [2] have been adopted to study the motion of human vocal cords. From the microwave approaches, it has been reported that the motions of vocal cords, providing essential information associated with phonation, can be detected by electromagnetic pulsed radar [3] and continuous-wave radars [4], [5]. The first electromagnetic pulsed radar system, called the glottal electromagnetic micropower sensor (GEMS), was presented by a research team at the Lawrence Livermore National Laboratory [3]. In [6], a thorough description of a speech phonation mechanism and vocal tract excitation model are given. Numerous extended applications were reported based on GEMS signals, including speaker recognition [7], [8], speech enhancement [9], speech recognition [10], [11], and speech coding [12]. Using a different detection mechanism, the continuous-wave coherent speech radar was developed in [4] and [5] with special emphasis on low-cost fabrication. Similar continuous-wave coherent radars have also been seen in the detection of the human cardiopulmonary signal, where the heartbeat rate is to be precisely detected from a remote distance [13]–[18]. The difference is that the speech radars concern the waveform quality of the vocal vibration signal, while the heartbeat radars focus on precise heartbeat rate detection.

The works of [3]–[5] achieved successful vocal vibration signal detection, but the detailed theoretical analysis of signal detection quality and the associated system circuit design are not given. With envisioning the great potential of vocal vibration detection in speech applications, this paper is aimed to a complete description of signal detection and system circuit design of the speech radar. The detection principle and the distortion of signal demodulation are described in Section II. The system circuit design is given in Section III. The experimental validation of developed speech radar circuitry, and experiments on vocal vibration detection of phonating vowels and words, are presented in Section IV. Finally, a conclusion is drawn in Section V.

II. Detection of Vocal Vibration Signal

A. Coherent Homodyne Demodulation

The human vocal vibration, as well as the cardiopulmonary movement, is in a periodic motion, which can be detected based on the Doppler radar principle. The cardiopulmonary signal detection focuses on the accurate detection of heartbeat and respiration rates [15]–[18], while the vocal vibration detection puts emphasis on the recovery of the original vocal excitation signal with minimum distortion. The single-path homodyne demodulation is adopted for the vocal vibration signal detection. The periodic opening and closing of vocal cords induces a phase modulation on the impinging electromagnetic wave. To correctly recover the phase-modulating signal, the coherent homodyne transceiver is used, as illustrated in Fig. 1, where $x(t)$ denotes the vocal vibration displacement and d is the averaged distance between the antenna and the vocal cords. The transmitted signal $T(t)$ is generated from a phase-locked local oscillator, which is amplified and radiated by the antenna toward the throat of human subject. A portion of $T(t)$ is fed into the receiver to coherently demodulate the received signal $R(t)$.

The received signal $R(t)$, intercepted by the receiver antenna, contains three components: the desired phase-modulated signal

Manuscript received February 12, 2010; accepted April 20, 2010. Date of publication July 15, 2010; date of current version August 13, 2010. This work was supported in part by the National Science Council, Taiwan.

C.-S. Lin was with the Department of Electrical Engineering, National Chung Cheng University, Chiayi 62102, Taiwan. He is now with MIPRO Electronics, Chiayi 600, Taiwan (e-mail: d93415013@ccu.edu.tw; jason@mipro.com.tw).

S.-F. Chang and C.-C. Chang are with the Department of Electrical Engineering and the Department of Communications Engineering, Center for Telecommunication Research, National Chung Cheng University, Chiayi 62102, Taiwan (e-mail: ieesfc@ccu.edu.tw; ccchang@ee.ccu.edu.tw).

C.-C. Lin is with the Department of Electrical Engineering, National Chung Cheng University, Chiayi 62102, Taiwan (e-mail: d95415002@ccu.edu.tw).

Digital Object Identifier 10.1109/TMTT.2010.2052968

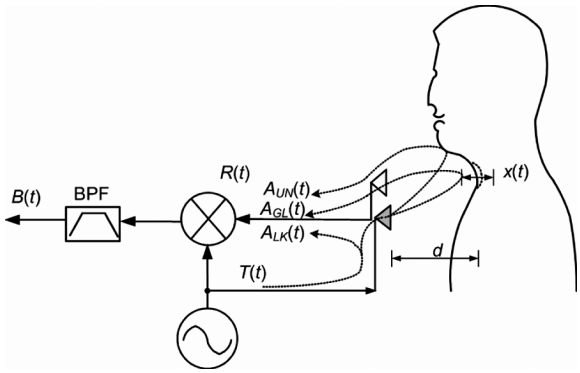

Fig. 1. Block diagram of the speech radar system.

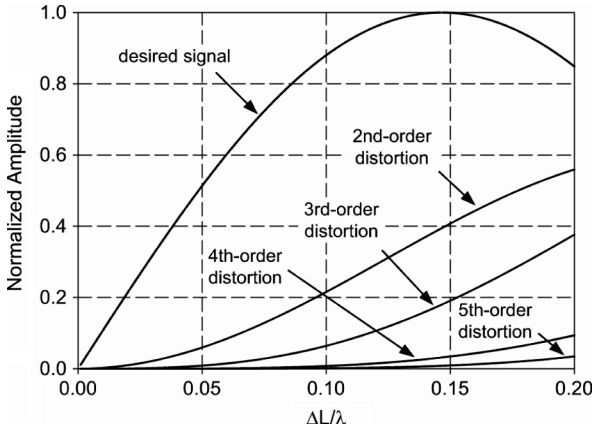

Fig. 2. Calculated harmonics of the demodulated signal for the single-tone vocal vibration excitation $x(t) = \Delta L \sin(\omega_{GL} t)$ for the case of $\phi_{\rm err} = 36°$.

$A_{GL}(t)$, the undesired interference $A_{\rm UN}(t)$, and the leakage component $A_{\rm LK}(t)$ from the transmitter. The $A_{GL}(t)$ is the backscattered signal from the vibrating vocal cords, which can be written as $V_{GL}\cos(\omega_c t + 4\pi x(t)/\lambda + \phi_{GL})$, where ω_c is the angular frequency of the transmitted electromagnetic wave, λ is the wavelength inside the human body, V_{GL} denotes the received signal amplitude, and ϕ_{GL} represents the round-trip phase difference between the received signal and the down-converting local oscillation signal. $A_{\rm UN}(t)$ represents the reflected signals from surrounding stationary organs and $A_{\rm LK}(t)$ denotes the leakage from the transmitter due to finite circuit isolation. Thus, $A_{\rm UN}(t) = V_{\rm UN}\cos(\omega_c t + \phi_{\rm UN})$ and $A_{\rm LK}(t) = V_{\rm LK}\cos(\omega_c t + \phi_{\rm LK})$. Since the same local oscillation signal is used for transmission and reception, the phase noise superimposed in the transmitted signal is carried over into the received components. Therefore, the down-converted baseband signal can be expressed as

$$\begin{aligned} B(t) =& G_{\rm RX}V_{GL}\cos\left(4\pi\frac{x(t)}{\lambda} + \phi_{GL} + \theta_{\rm PN}\right) \\ &+ G_{\rm RX}V_{\rm UN}\cos(\phi_{\rm UN} + \theta'_{\rm PN}) \\ &+ G_{\rm RX}V_{\rm LK}\cos(\phi_{\rm LK} + \theta''_{\rm PN}). \end{aligned} \tag{1}$$

In (1), $G_{\rm RX}$ denotes the total receiver gain. Since the same local oscillator is used for transmission and reception, the residual phase noises ($\theta_{\rm PN}$, $\theta'_{\rm PN}$, and $\theta''_{\rm PN}$) are dramatically suppressed [15]. The residual phase noise can then be neglected in the following distortion analysis. By the rejection from the bandpass filter in baseband, the near-dc components of down-converted $A_{\rm UN}(t)$ and $A_{\rm LK}(t)$ become negligible. Consequently, the normalized baseband signal becomes

$$\hat{B}(t) = \cos\left[4\pi\frac{x(t)}{\lambda} + \phi_{GL}\right]. \tag{2}$$

B. Distortion of Extracted Vocal Vibration Signal

1) Single-Tone Voiced Speech Excitation: Assuming that a single-tone voiced speech is excited, i.e., $x(t) = \Delta L\sin(\omega_{GL} t)$, where ΔL is the amplitude of vocal vibration signal and ω_{GL} denotes the angular vibrating frequency of vocal cords. Equation (2) then becomes

$$\hat{B}(t) = \cos\left(4\pi\frac{\Delta L}{\lambda}\sin\omega_{GL}t + \phi_{GL}\right). \tag{3}$$

By applying the Fourier expansion, $\hat{B}(t)$ can be decomposed into a summation of dc term and harmonics as follows:

$$\begin{aligned} \hat{B}(t) =& \,{\rm Re}\left\{e^{j[4\pi\Delta L\sin(\omega_{GL}t)/\lambda + \phi_{GL}]}\right\} \\ =& J_0\left(4\pi\frac{\Delta L}{\lambda}\right)\cos\phi_{GL} \\ &- 2\sum_{k=1}^{\infty} J_{2k-1}\left(4\pi\frac{\Delta L}{\lambda}\right)\sin\phi_{GL}\sin\left[(2k-1)\omega_{GL}t\right] \\ &+ 2\sum_{k=1}^{\infty} J_{2k}\left(4\pi\frac{\Delta L}{\lambda}\right)\cos\phi_{GL}\cos(2k\omega_{GL}t) \end{aligned} \tag{4}$$

where $J_k(x)$ is the kth-order Bessel function of the first kind. From (4), it is seen that the original excitation signal $x(t)$ is maximally recovered when ϕ_{GL} is odd multiples of $\pi/2$, but vanished when ϕ_{GL} equals to even multiples of $\pi/2$. Therefore, a phase shifter is inserted into the receiver local oscillator path to obtain maximal recovered signal intensity in the system implementation.

Even though ϕ_{GL} is tuned to an odd multiple of $\pi/2$, the human body motion can cause a phase tolerance to ϕ_{GL}. For example, an unintended random body motion of 1 mm induces a 2.4° tolerance at 1 GHz and an intended 15-mm body motion causes a 36° phase tolerance. This phenomenon gets worse when higher operation frequency is adopted. At 15 GHz, even the unintended random motion of 1 mm will cause a dramatic 36° tolerance. Therefore, in the following calculation, the phase tolerance $\phi_{\rm err}$ is included in the optimal ϕ_{GL}. The first five harmonics of the demodulated baseband signal for the case of $\phi_{\rm err} = 36°$ are shown in Fig. 2. The desired signal intensity increases with respect to the normalized vocal cords displacement $\Delta L/\lambda$ and peaks around $\Delta L/\lambda = 0.15$. However, the higher order harmonics increase as well. Therefore, the total harmonic

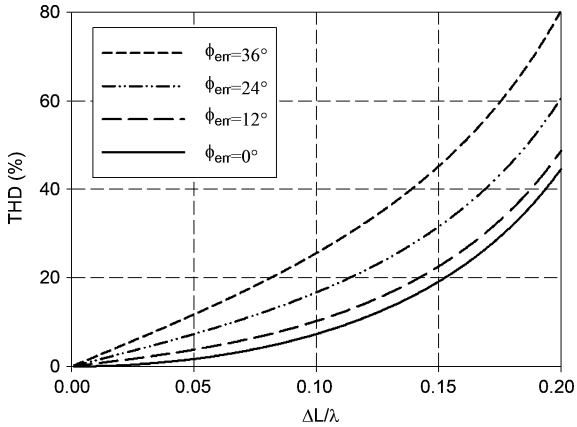

Fig. 3. THD versus the normalized displacement $\Delta L/\lambda$ for a single-tone voice excitation with different random-motion phase tolerances.

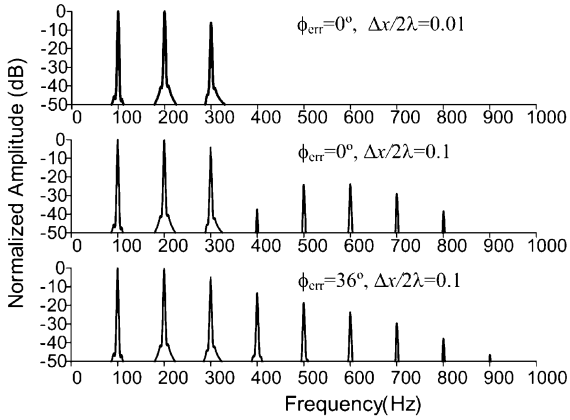

Fig. 4. Simulated results of the demodulated baseband spectra for a three-tone excitation.

distortion (THD) is calculated for the signal quality assessment, defined in (5) at the bottom of this page.

Fig. 3 shows the calculated THD with respect to the normalized displacement $\Delta L/\lambda$ for $\phi_{\rm err} = 0°$, 12°, 24°, and 36°, respectively. The THD increases with $\Delta L/\lambda$ and also with the body random-motion phase tolerance $\phi_{\rm err}$. Fig. 3 gives a guideline for selecting the operation frequency of the speech radar system. For a design example, the vocal cords displacement ΔL is around 0.5 mm, the random body motion is 1 mm, and the relative dielectric constant of the interior neck is 40 [19]. If the specified THD is less than 0.15%, the operation frequency must be selected at less than 1 GHz.

2) Multitone Voiced Speech Excitation: In general, the vocal cords displacement $x(t)$ is nonsinusoidal and is more appropriately expressed as a summation of N sinusoidal tones

$$x(t) = \sum_{n=1}^{N} \Delta L_n \sin\left(n\omega_{GL}t + \theta_n\right) \tag{6}$$

where $\Delta L_n \sin(n\omega_{GL}t + \theta_n)$ is the nth harmonic of the vocal vibration displacement $x(t)$. Therefore, the demodulated baseband signal becomes (7), shown at the bottom of this page.

It can be seen from (7) that the nonlinearity of the coherent vocal signal demodulation causes vast intermodulation harmonics. Consider a voiced speech model, which has three in-phase tones with normalized amplitudes of 0 dB at 100 Hz, 0 dB at 200 Hz, and -6 dB at 300 Hz. The calculated spectra of baseband signal are illustrated in Fig. 4, where Δx is defined as $x(t)_{\max} - x(t)_{\min}$. When the normalized displacement of $\Delta x/2\lambda$ is 0.01 and $\phi_{\rm err} = 0°$, the demodulated signal has three primary tones at 100, 200, and 300 Hz, respectively, and all other intermodulated harmonics are smaller than -50 dB. However, when $\Delta x/2\lambda$ is increased to 0.1, the intermodulation distortion is degraded to -24 dB for $\phi_{\rm err} = 0°$ and even worse to -14 dB for $\phi_{\rm err} = 36°$.

III. System Circuit Design

A. System Description

The speech radar was realized with the circuit block diagrams in Fig. 5. A phase-locked-loop frequency synthesizer generates a stable continuous-wave signal, where part of it is fed back through a phase shifter to the receiver and the other part, filtered by a low-pass filter for harmonic suppression, is radiated by the transmitter antenna. As the transmitted wave impinges the vibrating glottis, a phase-modulated signal is back-scattered

$$\text{THD} = \frac{\text{root-mean-squared summantion of harmonics}}{\text{root-mean-squared value of fundamental harmonic}} = \frac{\sqrt{\sum_{k=2}^{\infty}\left[J_{2k-1}\left(4\pi\frac{\Delta L}{\lambda}\right)\sin\phi_{GL}\right]^2 + \sum_{k=1}^{\infty}\left[J_{2k}\left(4\pi\frac{\Delta L}{\lambda}\right)\cos\phi_{GL}\right]^2}}{J_1\left(4\pi\frac{\Delta L}{\lambda}\right)\left|\sin\phi_{GL}\right|} \tag{5}$$

$$\hat{B}(t) = \cos\left[4\pi\sum_{n=1}^{N}\frac{\Delta L_n}{\lambda}\sin\left(n\omega_{GL}t + \theta_n\right) + \phi_{GL}\right] = \text{Re}\left\{e^{j\phi_{GL}}\prod_{n=1}^{N}\left[\sum_{k_n=-\infty}^{\infty}J_{k_n}\left(4\pi\frac{\Delta L_n}{\lambda}\right)e^{jk_n(n\omega_{GL}t+\theta_n)}\right]\right\} \tag{7}$$

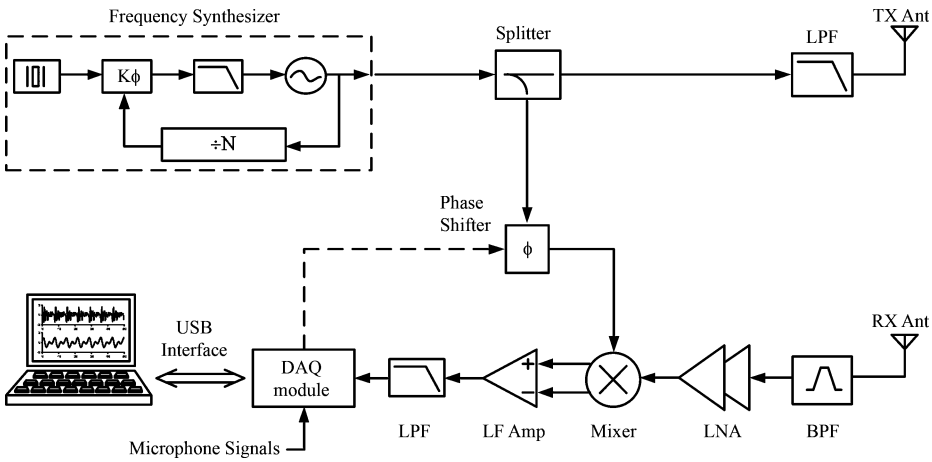

Fig. 5. Block diagram of the speech radar system.

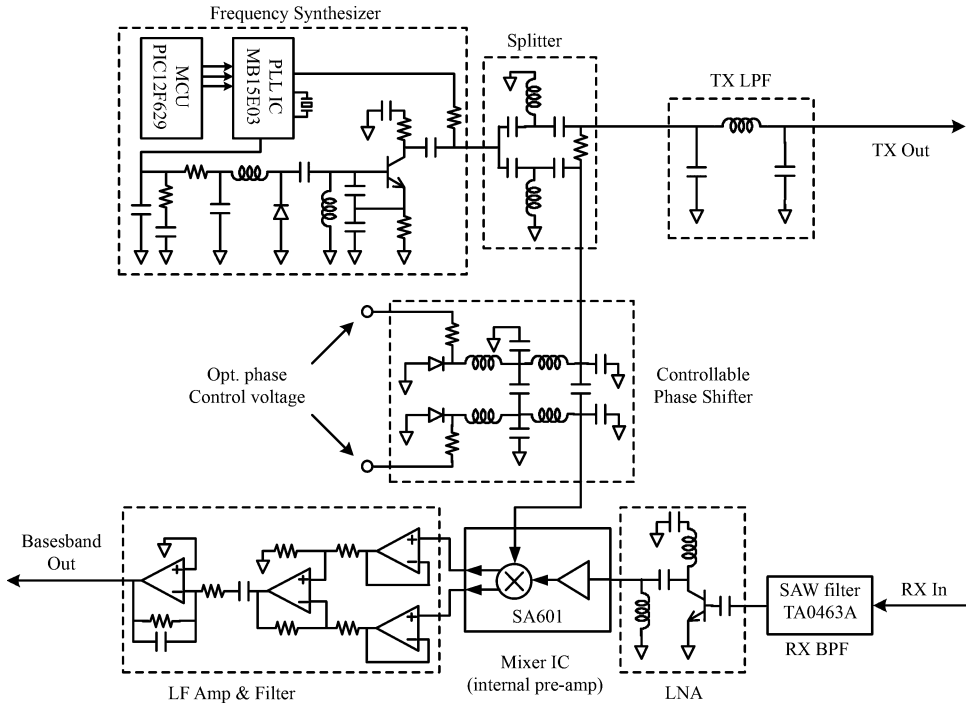

Fig. 6. Schematic circuit diagrams of the key building blocks.

and intercepted by the receiver antenna. The reflections from other organs can be picked up as well, depending on the antenna pattern.

The received signals, including the desired phase-modulated signal and other unwanted interferences, are filtered by a band-pass filter to eliminate out-of-band interferences. The filtered signal is then amplified by the low-noise amplifier and coherently down-converted into the baseband. In the baseband domain, a low-frequency amplifier, followed by an antialiasing filter, optimizes the signal-to-noise ratio for driving the followed data acquisition (DAQ) module. The DAQ samples the baseband signal for executing data computation in the digital domain. The DAQ module also controls the phase shifter to meet the $\phi_{GL} = \pi/2$ requirement for obtaining a maximal baseband signal intensity, as discussed in Section II-B. In addition to the speech radar system, an acoustic microphone system is also incorporated to receive the acoustic speech, which is used for comparison with the electromagnetic vocal vibration signal.

B. Circuit Design

The schematic circuit diagrams of key building blocks are shown in Fig. 6. According to the analysis in Section II, the operating frequency is selected at 925 MHz. The circuit details are described as follows.

1) *Frequency synthesizer*: A phase-locked-loop integrated-circuit MB15E03 and microcontrol unit PIC12F629 are integrated with a 925-MHz Colpitts voltage-controlled oscillator to realize the frequency synthesizer. The synthesized frequency range is from 905 to 945 MHz in a 25-kHz step. The synthesized microwave signal has a 2-mW output power with phase noise of −120 dBc/Hz at 100-kHz offset from the carrier.

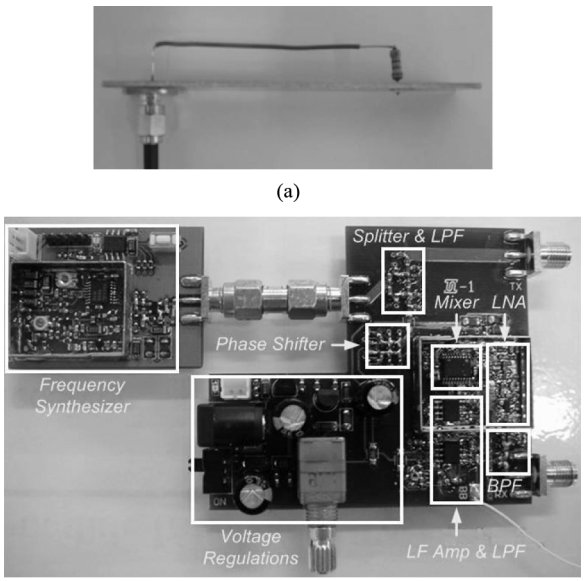

Fig. 7. Photograph of the developed speech radar system. (a) Antenna. (b) Transceiver circuits.

2) *Power splitter*: The power splitter was realized by a lumped T topology, which has a measured insertion loss smaller than 1 dB from 905 to 945 MHz. The amplitude imbalance and phase imbalance are within 0.5 dB and 5°, respectively. The output isolation is greater than 20 dB.
3) *Transmitter low-pass filter*: The transmitter low-pass filter was implemented by a third-order Butterworth topology. The circuit has an insertion loss less than 1 dB in 905–945 MHz and a stopband rejection greater than 20 dB in 1800–1900 MHz.
4) *Receiver bandpass filter*: A surface-acoustic-wave filter, centered at 925 MHz, was used as the receiver filter. The 1-dB passband is 6 MHz and the insertion loss is 4.2 dB, with stopband rejection better than 36 dB at 40 MHz away from the passband center.
5) *Low-noise amplifier and mixer*: The low-noise amplifier was designed with a common-emitter bipolar transistor, which has a measured gain of 15 dB, noise figure of 1.6 dB, and input $P_{1\mathrm{dB}}$ of 2 dBm. The SA601 was used as a coherent down-mixer, which contains a pre-amplifier and a wide dynamic-range mixer. The pre-amplifier has an 11-dB gain, 1.6-dB noise figure, and 0-dBm input third-order intercept point (IP3). The mixer has a conversion gain of 6.5 dB, noise figure of 9.5 dB, and input IP3 of -2 dBm.
6) *Phase shifter*: A reflection-type phase shifter was designed for the phase tuning of local oscillation signal [20]. It includes a lumped-element branch-line coupler, terminated with two series-resonant surface-mounted silicon hyperabrupt-junction varactor diodes. The phase tuning range is larger than 190° with a 2-dB insertion loss.
7) *Low-frequency amplifier and anti-aliasing lowpass filter*: Operational amplifiers are used to realize the low-frequency gain amplifier and active low-pass filter. The low-pass filter has a cutoff frequency of 4 kHz to reject unwanted interference, which acts as the antialiasing filter for the followed data-acquisition module.
8) *DAQ module*: The National Instrument DAQ module USB-9233 was used for baseband vocal vibration signal acquisition. It has four simultaneous process channels with 24-bit resolution, 102-dB dynamic range, and a maximum sampling rate of 50 kS/s.
9) *Antennas*: The beverage (wire) antenna at 925 MHz was designed as a near-field induction antenna.

The photograph of the developed speech radar system is illustrated in Fig. 7.

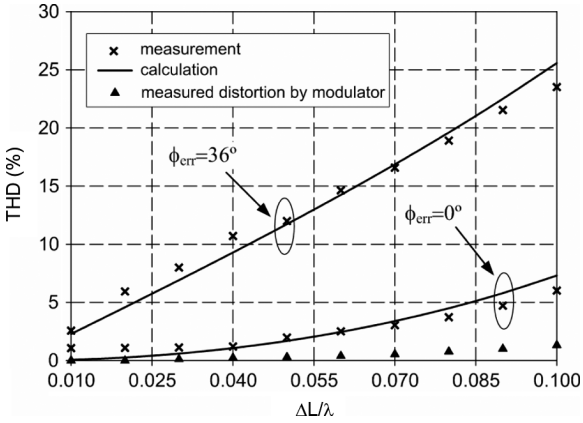

Fig. 8. Measured and calculated THD for a 100-Hz single-tone excitation.

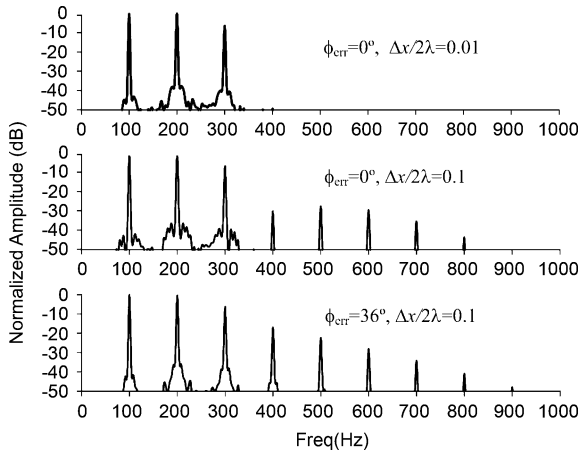

Fig. 9. Measured results of the demodulated baseband spectra for a three-tone excitation.

Fig. 10. Experimental setup of the detection of vocal vibration signal and acoustic signal. (a) Targeted person with microwave antennas and acoustic microphone. (b) Speech radar circuit.

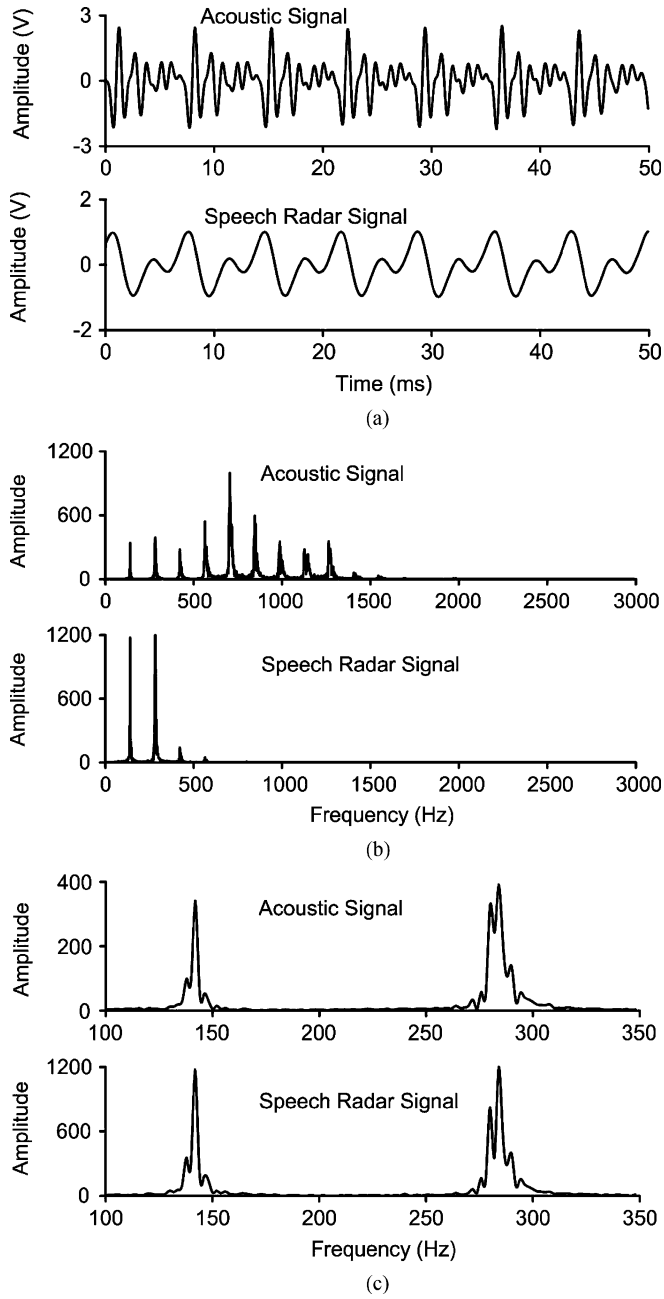

Fig. 11. Measurement waveforms of the vowel /a/ by the acoustic microphone system (acoustic signal) and the speech radar system (vocal vibration signal). (a) Time-domain waveforms. (b) Frequency-domain spectra. (c) Detailed spectral waveform of the fundamental and second harmonics.

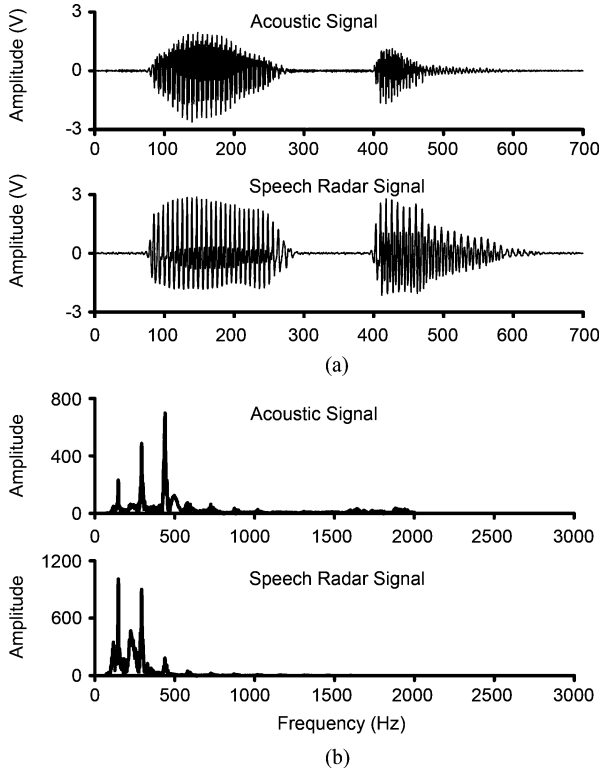

Fig. 12. Measurement waveforms of the word "Jason" by the acoustic microphone system (acoustic signal) and the speech radar system (vocal vibration signal). (a) Time-domain waveforms. (b) Frequency-domain spectra.

IV. Experimental Results

A. Experimental Validation of Implemented Coherent Speech Receiver

Before conducting the voiced speech detection of human subjects, the implemented speech receiver was validated experimentally. An analog voltage-controlled phase shifter [20] was used as a phase modulator to produce the desired phase-modulated electromagnetic signal. The modulator distortion was measured and illustrated in Fig. 8, which indicates that the modulation distortion is less than 1.3% for $\Delta L/\lambda$ up to 0.1. This ensures the modulator distortion is kept well below the demodulation distortion.

First, a 925-MHz signal, phase modulated by a 100-Hz sinusoidal waveform, was generated from the above phase modulator and fed into the implemented speech receiver. The THD of the output demodulation signal is illustrated in Fig. 8. The measured THD is within 1.1%–6.0% for $\phi_{\rm err} = 0^\circ$ and 2.6%-23.5% for $\phi_{\rm err} = 36^\circ$ when $\Delta L/\lambda$ is in the range of 0.01–0.1. The measured results agree very well with the calculation from (5).

Second, another 925-MHz signal is modulated by a three-tone waveform, containing 100-, 200-, and 300-Hz sinusoidal frequency with normalized amplitudes of 0, 0, and −6 dB, respectively. The measured spectra are plotted in Fig. 9. As expected, serious intermodulation products appear at $\Delta x/2\lambda = 0.1$ for the case of $\phi_{\rm err} = 36^\circ$. This result is consistent with the simulation in Fig. 4. These experimental results validate the implemented speech radar circuits.

B. Vocal Vibration Signal Detection

The experiment setup for the detection of vocal vibration signal and acoustic signal is illustrated in Fig. 10. For the speech radar system, antennas are lightly placed on the larynx area of the human subject. Meanwhile, the acoustic signal is also recorded simultaneously by the microphone for data comparison. A 30-year-old male participant phonates in a variety of vowels and words.

The experiment on the vowel of /a/ was conducted. The time-domain waveforms of the speech radar signal and acoustic

signal are displayed in Fig. 11(a), which indicate the same fundamental period, but are quite different on the waveform shape. Fig. 11(b) and (c) further illuminate their similarities and differences in the frequency domain. The fundamental and second harmonics detected in the speech radar signal are consistent with those in the acoustic signal, while extra higher harmonics exist in the acoustic signal. This result can be understood from the voiced speech articulation. When generating voiced speech, the vocal cords are pushed open by the air pressure from the lungs and then close because of the folds' natural elasticity [21]. Therefore the vocal cords vibrate at the fundamental frequency with few harmonics, from which the waveform is detected by the speech radar system. Instantly, with the compressed air flow to the super-glottis, it is modulated by the vocal tract, tongue, lips, and jaw, generating rich harmonics in the spoken acoustic waveform.

A further experiment was conducted on word detection. The word "Jason" was pronounced and the detected waveforms are shown in Fig. 12. The time-domain waveforms from the speech radar system have a consistent segment pattern when compared with those from the acoustic microphone system. From the aspect of the frequency domain, the fundamental and second harmonics of the speech radar signal are at the same frequencies with those from the acoustic signal, but they have different relative amplitudes due to the modulation of the vocal tract, tongue, lips, and jaw.

V. Conclusion

In this paper, the thorough analysis and system design of the speech radar is presented. To be able to detect the small phase fluctuation due to the tiny vocal vibration displacement, typically in millimeters, a coherent homodyne transceiver is used,where the phase noise can be significantly suppressed by the phase coherence of the transmitted and received signals. The THD is derived to assess the detected vocal vibration signal quality, where the single-tone and multitone models for the vocal cords waveform are considered. The effect of intended and unintended body motion is also included. From the derived result, the operation frequency of speech radar can be properly determined.

Experiments on the detection of vowel and word phonations were conducted, while the conventional acoustic microphone detection was performed, as well for the purpose of comparison. The results show that the measured speech radar signals have excellent consistency with the acoustic signals, which validates the speech detection capability of the proposed radar system. The vocal vibration waveforms were found to have power content distribution dominantly over the first few harmonics, which is useful to the estimation of the sound pitch. In contrast, the acoustic waveforms contain richer harmonics, generated from the modulation by the vocal organs such as tract, lips, and jaw.

The speech radar is essentially immune to background acoustic noise, making it more appealing for applications in high background acoustic-noise environments where acoustic signals are inaccessible or blocked. It shows a variety of potential applications, including the background acoustic noise removal, speaker verification, and identification, as well as medical uses.

References

[1] D. G. Childers and J. N. Larar, "Electroglottography for Laryngeal function assessment and speech analysis," *IEEE Trans. Biomed. Eng.*, vol. BME-31, no. 12, pp. 807–817, Dec. 1984.

[2] T. Wittenberg, P. Mergell, M. Tigges, and U. Eysholdt, "Quantitative characterization of functional voice disorders using motion analysis of high-speed video and modeling, acoustics, speech, and signal processing," in *IEEE Int. Acous., Speech, Signal Process. Conf.*, Apr. 1997, vol. . 3, pp. 1663–1666.

[3] J. F. Holzrichter, G. C. Burnett, L. C. Ng, and W. A. Lea, "Speech articulator measurements using low power EM wave sensor," *J. Acoust. Soc. Amer.*, vol. 103, no. 1, pp. 622–622, 1998.

[4] C. Liang, "Novel microwave and millimeter-wave radar technologies and applications," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California at Davis, Davis, CA, 2001.

[5] Y. Yang, "Radar based speech analysis system," Master's thesis, Dept. Elect. Comput. Eng., Univ. California at Davis, Davis, CA, 2000.

[6] G. C. Burnett, "The physiological basis of glottal electromagnetic micropower sensor (GEMS) and their use in defining an excitation function for the human vocal tract," Ph.D. dissertation, Dept. Appl. Sci., Univ. California at Davis, Davis, CA, 1999.

[7] T. J. Gable, "Speaker verification using acoustic and glottal electromagnetic micropower sensor (GEMS) data," Ph.D. dissertation, Dept. Appl. Sci., Univ. California at Davis, Davis, CA, 2000.

[8] W. M. Campbell, T. F. Quatieri, J. P. Campbell, and C. J. Weinstein, "Multimodal speaker authentication using nonacoustic sensors," in *Proc. Multimodal User Authenatication Workshop*, 2003, pp. 215–222.

[9] L. C. Ng, G. C. Burnett, J. F. Holzrichter, and T. J. Gable, "Denoising of human speech using combined acoustic and EM sensor signal processing," *Proc. Int. Acoust., Speech, Signal Process. Conf.*, vol. 1, pp. 229–232, 2000.

[10] B. Raj and R. Singh, "Feature compensation with secondary sensor measurements for robust speech recognition," in *Proc. EUSIPCO*, 2005, 4 pp.

[11] C. Demiroglu and D. V. Anderson, "Broad phoneme recognition in noisy environments using the GEMS device," in *Proc. Asilomar Signals, Syst. Comput. Conf.*, 2004, vol. 2, pp. 1805–1808.

[12] T. F. Quatieri, K. Brady, D. Messing, J. P. Campbell, W. M. Campbell, M. S. Brandstein, C. J. Clifford, J. D. Tardelli, and P. D. Gatewood, "Exploiting nonacoustic sensors for speech encoding," *IEEE Trans. Audio Speech Language Process.*, vol. 14, no. 2, pp. 533–544, Mar. 2006.

[13] K. M. Chen, D. Misra, H. Wang, H. R. Chuang, and E. Postow, "An X-band microwave life-detection system," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 7, pp. 697–702, Jul. 1986.

[14] K. M. Chen, Y. Huang, J. Zhang, and A. Norman, "Microwave life-detection systems for searching human subjects under earthquake rubble and behind barrier," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 1, pp. 105–114, Jan. 2000.

[15] A. D. Droitcour, O. Boric-Lubecke, V. M. Lubecke, J. Lin, and G. T. A. Kovac, "Range correlation and I/Q performance benefits in single-chip silicon Doppler radars for noncontact cardiopulmonary monitoring," *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 3, pp. 838–848, Mar. 2004.

[16] Y. Xiao, J. Lin, O. Boric-Lubecke, and V. M. Lubecke, "Frequency tuning technique for remote detection of heartbeat and respiration using low-power double-sideband transmission in Ka-band," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 5, pp. 2023–2032, May 2006.

[17] B. Park, O. Boric-Lubecke, and V. M. Lubecke, "Arctangent demodulation with DC offset compensation in quadrature Doppler radar receiver systems," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 5, pp. 1073–1079, May 2007.

[18] C. Li and J. Lin, "Random body movement cancellation in Doppler radar vital sign detection," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 12, pp. 3143–3152, Dec. 2008.

[19] T. Yilmaz, T. Karacolak, and E. Topsakal, "Characterization and test of a skin mimicking material for implantable antennas operating at ISM band (2.4 GHz–2.48 GHz)," *IEEE Antennas Wireless Propag. Lett.*, vol. 7, no. 11, pp. 418–420, Nov. 2008.

[20] C. S. Lin, S. F. Chang, C. C. Chang, and Y. H. Shu, "Design of a reflection-type phase shifter with wide relative phase shift and constant insertion loss," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 9, pp. 1862–1868, Sep. 2007.
[21] C. Rowden, *Speech Processing*. New York: McGraw-Hill, 1992.

Chien-San Lin was born in Chiayi, Taiwan, in 1976. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Chung Cheng University, Taiwan, in 1996, 2000, and 2010, respectively.

In 2003, he joined MIPRO Electronics, Chiayi, Taiwan, where he is currently a Research and Development Engineer engaged in the development of professional microphone systems. His research interests include RF circuit designs in wireless communication systems.

Sheng-Fuh Chang (S'83–M'84–SM'07) received the B.S. and M.S. degrees in communications engineering from National Chiao-Tung University, Taiwan, in 1982 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Wisconsin–Madison, in 1991.

He was involved with high-power microwave and millimeter-wave sources such as free-electron lasers and Cerenkov masers with the Center for Plasma Theory and Computation, University of Wisconsin–Madison. In 1992, he joined the Hyton Technology Corporation, where he was responsible for C- and Ku-band satellite low-noise down-converter and multichannel multipoint distribution system (MMDS) transceivers. In 1994, he joined the Department of Electrical Engineering, National Chung Cheng University, Taiwan, where he is currently a Full Professor with the Department of Electrical Engineering and Vice Director of the Center for Telecommunication Research. His research interests include microwave and millimeter-wave integrated circuits (ICs) with CMOS, HBT, and pseudomorphic HEMT (pHEMT) technologies, multifunctional RF transceivers, smart-antenna RF system, and high-power microwave sources.

Prof. Chang is a member of Phi Tau Phi and Sigma Xi.

Chia-Chan Chang (S'99–M'04) was born in Tainan, Taiwan, in 1973. She received the B.S. degree in communication engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 1995, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Davis (UCD), in 2001 and 2003, respectively.

From 1995 to 1997, she was a Full-Time Teaching Assistant with the Department of Electronics Engineering, National Chiao-Tung University. Since February 2004, she has become an Assistant Professor with the Department of Electrical Engineering, National Chung-Cheng University, Chiayi, Taiwan. She also holds a joint-appointment with the Department of Communications Engineering, National Chung-Cheng University. Her research currently focuses on phased antenna array technology developments, microwave/millimeter-wave circuit designs, and the application of radar systems.

Chun-Chi Lin (S'06) was born in Chia-Yi, Taiwan, in 1979. He received the B.S. degree in electrical engineering from Yuan Ze University, Taoyuan, Taiwan, in 2002, the M.S. degree in electrical engineering from Chung-Cheng University, , Chiayi, Taiwan, in 2006, and is currently working toward the Ph.D. degree in electrical engineering at Chung-Cheng University.

His primary research interests are 3-D human electromagnetic (EM) models, coated antennas, chip antennas, and chip antenna measurement systems. His additional research interests include antenna arrays and RF/microwave circuits for wireless applications.