Sensors and Actuators A 191 (2013) 99–104

Contents lists available at SciVerse ScienceDirect

Sensors and Actuators A: Physical

journal homepage: www.elsevier.com/locate/sna

Smart radar sensor for speech detection and enhancement

Ying Tian [1], Sheng Li [1], Hao Lv, Jianqi Wang *, Xijing Jing

College of Biomedical Engineering, The Fourth Military Medical University, Xi'an 710032, Shaanxi, PR China

ARTICLE INFO

Article history:
Received 11 July 2012
Received in revised form 3 December 2012
Accepted 4 December 2012
Available online xxx

Keywords:
Radar sensor
Speech detection
Third-order cumulant
Bispectrum
Speech enhancement

ABSTRACT

Speech communication with microphone requires the user to wear additional gears or have a set of kits placed at a short distance from the speaker. Such gears limit the user' activity and cause inconvenience in certain condition. This study attempts to use a new radar sensor as a non-contact method for speech detection. To achieve this kind of non-contact detection, a smart radar sensor was developed to capture the small vibration from the larynx of a speaker. This radar sensor operates at 35.5 GHz and uses two parabolic antennas. A superheterodyne receiver is also employed to reduce the DC offsets and $1/f$ noise. This radar sensor provides a new method for speech communication. However, more complex noise components present in the output speech signal. Harmonics of millimeter wave and electrocircuit noise are the main part of the noise source. Therefore, to reduce the noise introduced from the radar sensor, we propose a higher-order statistics (HOS) algorithm suitable for the radar speech. First, third-order cumulant is applied to identify voiced frame from unvoiced one, and then bispectrum algorithm is used to further reduce the noise because of its superiority in suppressing additive Gaussian noise. The performance of the algorithm is evaluated both by means of subjective and objective measures. Results demonstrate that the microwave radar sensor can successfully detect speech signal from the speaker. In addition, the speech quality is improved significantly by the proposed algorithm. The smart radar sensor meets the requirement of non-contact detection. It will bring new light on free communication and extend the application of radar to speech detection.

1. Introduction

The traditional method for speech communication is through the microphone. The most commonly used are acoustic microphones. Those microphones use various material and techniques to produce an electrical voltage signal from mechanical vibration [1–3]. For example, Pedersen et al. fabricated silicon-based condenser microphones with polyimide diaphragm and backplate [4]. Kronast et al. designed a single-chip condenser microphone using porous silicon as sacrificial layer for the air gap [5]. The other type of microphone is non-acoustic microphone. For example, physiological microphone (PMIC) is a non-acoustic sensor which captures speech via skin vibrations through contact with the skin near the cricoid and thyroid cartilage [6]. Throat microphone is also used in automatic speech recognition on motorcycles. It is put around the neck with the transducer placed on the larynx with slight pressure. The throat microphone picks up the mechanical vibration of larynx to extract speech information [7]. Bone conduction microphone can be mounted on the user's head, and record/transmit signal responsive to the auditory vibrations [8]. However, these microphones need to be mounted on the human body through some specific gears such as a helmet or a belt or placed close to the speaker's mouth. Wearing those kits inevitably restricts the user' activity and causes discomfort and tension. Some microphones even require careful placement and a small deviation may lead to a failure of the measurement.

Therefore we propose a new non-contact method for speech communication by using the microwave radar sensor. This method does not need the speaker to wear any additional gears or have such gears placed close in front of them, and thus is more convenient and comfortable for its user. The principle of speech detection by radar sensor is that radar radiates electromagnetic waves to the larynx of the human body, and receives echo waves, which are modulated by the vibration of skin around the larynx. Thus the speech information can be extracted according to the phase variations of the echo waves [9,10]. Some previous research also verifies the feasibility of our study. As early as 1996, Li's group successfully detected the speech signal with a 40 GHz Doppler radar, and the experimental results confirm the feasibility of using millimeter wave (MMW) Doppler radar for speech detection [11]. Later in 1998, Holzrichter et al. discussed the application of low power electromagnetic (EM) wave sensors in the measurement of speech articulator motions

* Corresponding author. Tel.: +86 29 84774843; fax: +86 29 84774843.
E-mail address: cyt_bora@sina.com (J. Wang).
[1] These authors contributed equally to this work and should be regarded as co-first authors.

0924-4247/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sna.2012.12.002

[12]. In 2005, Hu and Raj attempted to use the reflected signal of acoustic Doppler radar from a person' mouth to determine if the person is speaking or not [13]. Those studies can extract small vibration information from the radar echo. However, the previous work concentrated on the relationship between organ motion and voice activity, seldom dedicating to speech communication.

A microwave radar sensor for non-contact speech detection is developed in our laboratory. This radar sensor operates at 35.5 GHz. The high operation frequency has excellent sensitivity for the acquisition of speech signal. It uses two parabolic antennas to enhance the stability of the system. A superheterodyne receiver is also employed to reduce the DC offsets and $1/f$ noise. During the detection experiment, it was observed that the echo signal was disturbed by unpleasant noise, which is especially true for the low-frequency components of the speech. In listening tests, the radar speech showed certain artificial quality. The degraded quality and intelligibility of the radar speech are mainly due to the presence of harmonics of millimeter wave and electrocircuit noise.

Due to the special characteristics possessed by the radar speech, it is required to apply appropriate algorithm on the output signal in order to extract the needed speech information from those noise components. The commonly used algorithms for speech enhancement include spectral subtraction and Wiener filtering [14]. While spectral subtraction reduces the broadband noise, it also usually introduces an annoying "musical noise" [15]. Regarding Wiener filtering, it does not consider the effect of the amplitude of the frequency component on human hearing and hence often causes serious speech distortion. Typically, the received wave form consists of the speech signal plus additive Gaussian receiver noise [16]. Hence in this paper we propose higher-order statistics algorithm to suppress the noise and improve the quality of the radar speech. We consider employing the special case of higher-order statistics: bispectrum. Bispectrum is more suitable for the handling of random signals. It contains richer information than first and second order statistics such as power spectrum. More importantly, bispectrum is superior in suppressing additive Gaussian noise, whereas the power spectrum contains noise power in all the frequency components [17].

This non-contact microwave radar sensor makes the user free from the limitation by using a microphone. In addition, it provides us a more convenient method for voice communication without additional gears contacting with human body, as well as bringing a new experience for free communication.

2. Description of the radar sensor

The block diagram of the radar sensor is shown in Fig. 1. The sensor includes two antennas: transmitting antenna and receiving antenna, and they deal with the transmission and receiving of the microwave signal respectively. The two antennas can reach a maximum gain of 38.5 dB at 35.5 GHz and their estimated beam width is 9°. The 35.5 GHz microwave signal is produced by mixing the volt control oscillator (VCO) signal with crystal oscillator (CO) signal and its power level is controlled by the variable attenuator (0–35 dB). When the speech signal is captured by the receiving antenna, it first passes through a low noise amplifier (LNA) and then is converted to a signal with much lower frequency components suitable for processing by, two down-converters (Mixer2 and Mixer3). After passing through the Mixer3, the signal is sent to a band-pass filter (BPF, frequency from 100 Hz to 5000 Hz, which is close to the frequency range of the human voice). Before the signal is processed by the computer, it is sampled via a 16-channel A/D converter. Then the final signal can be sent to a D/A converter and pass through a power amplifier (PA) to drive a speaker. The superheterodyne receiver is a property of the radar sensor. It is represented by the dashed box in Fig. 1. Such structure can reduce DC offsets and $1/f$ noise, so that the signal-to-noise ratio and the detection sensitivity for small high frequency speech signals are improved [18].

3. Methods

3.1. Experiments

Ten volunteers (five males and five females) with ages ranging from 20 to 40 years participated in the experiment. All of the subjects are native speakers of Mandarin Chinese and they reported no history of ear pathology. The experiment was conducted in a laboratory room that covers an area of 172.5 m^2 and the speech was recorded in relatively quiet environment. The speakers were required to sit still on a stool during the experiment, and the radar sensor was placed in front of each speaker at a distance of 10 m without any obstacles between the sensor and speakers. They were instructed to speak seven test sentences at normal loudness and speaking rate. The speech signal was captured by the radar sensor and recorded by a computer connected with the signal output of radar via a data line. Then, we used the MATLAB software

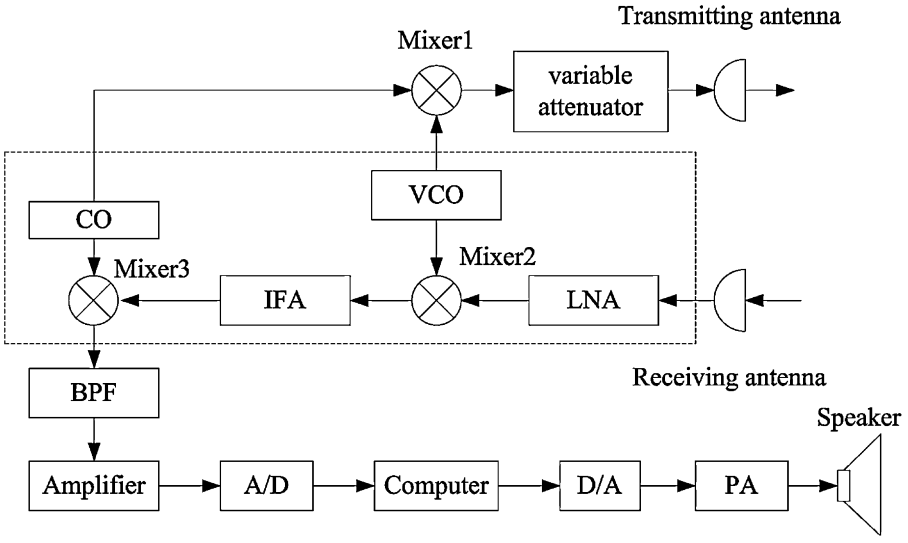

Fig. 1. Block diagram of the radar sensor.

package (MATLAB version 6.5; The Math Works, Inc.; Natic, MA, USA) for further processing [18].

3.2. Evaluations

To evaluate the performance of the proposed algorithm to microwave radar speech signal, Wiener filtering and spectral subtraction algorithm were also used for speech enhancement to make a comparison of effectiveness with our algorithm. Three types of noises, white noise, pink noise and babble noise, taken from the NOISEX-92 database, were added to the radar speech at four different SNR levels (0, 5, 10, 20 dB) respectively. These speech signals corrupted by the three noises were enhanced by the three algorithms respectively. Then both the subjective and objective measures were applied to analyze their performance.

The objective measure uses log spectral distance (LSD), given by

$$\mathrm{LSD} = \frac{1}{L}\sum_{l=1}^{L}\sqrt{\frac{1}{K/2+1}\sum_{k=0}^{K/2}(10\log_{10}|\hat{S}[k]| - 10\log_{10}|S[k]|)^2} \quad (1)$$

where k and K denote Discrete Fourier Transform index and the index of π rad, respectively. This measure evaluates how similar the log-spectrum of enhanced speech and the clean speech so that the smaller the LSD is, the closer the shape of log-spectrum of enhanced speech and the clean speech [19].

The subjective measure contains two methods. One is waveforms and spectrograms, and the other is listening tests. In the listening test, the 10 subjects were instructed to listen to recordings, which were collected from enhanced speech materials under pink noise, white noise and babble noise with SNR set at 0 dB, and rate the overall quality of played signals using a standard mean opinion score (MOS) scale, where MOS is a commonly used subjective test with a five-point quality scale ranging between one and five [20]. The final MOS results were obtained by averaging the scores over the 10 listeners for each algorithm.

4. The proposed algorithm for radar sensor speech

Here we consider two important concepts of HOS: cumulant and its corresponding higher-order spectra.

One significant property of cumulant is all the cumulants of Gaussian random processes are zero when the order is greater than two. So if a non-Gaussian signal is received along with additive Gaussian noise, a transform to a higher-order cumulant domain will eliminate the noise [21]. For that reason we apply HOS to radar speech signal.

In our algorithm, we focus on the third-order cumulant and bispectrum.

Assume $x(n)$, $n = 0, \pm1, \pm2, \ldots$ is a real stationary discrete-time sequence with zero mean. The third-order cumulant of $x(n)$ is defined:

$$c_{3x}(\tau_1, \tau_2) = E\{x(n)x(n+\tau_1)x(n+\tau_2)\} \quad (2)$$

and the Fourier transform of the third-order cumulant is called a bispectrum. Here the bispectrum of $x(n)$ is given by

$$B(\omega_1, \omega_2) = H(\omega_1)H(\omega_2)H^*(\omega_1+\omega_2) \quad (3)$$

where $H(\omega)$ is the Fourier transform of $x(n)$ [22].

The two Eqs. (2) and (3) form the basis of our algorithm.

The flowchart of our algorithm is shown in Fig. 2.

We segment our speech data into n records and each frame has a length of 128 samples with 50% of overlap. Each short-time frame is regarded as a stationary sequence for further processing. To estimate bispectrum of each frame, called sample bispectrum $B_i(\omega_1, \omega_2)$ [23], a direct method, called conventional periodogram type bispectrum estimator [24], is used in this paper.

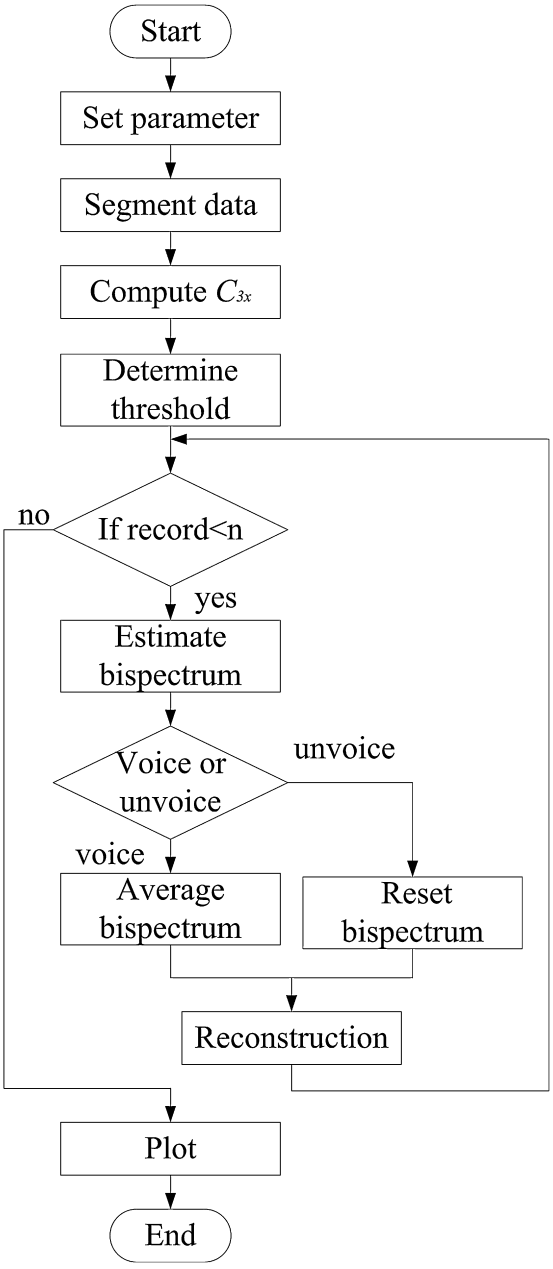

Fig. 2. Flow chart of the proposed algorithm for radar sensor speech.

In order to distinguish voiced frame from unvoiced one, the third-order cumulant is applied in this algorithm. This technique is based on the fact that unvoiced frame is produced by a Gaussian-like excitation and thus results in a small third-order cumulant whereas the condition is not true for voiced frame. Hence, we assign a threshold for third-order cumulant of each frame to determine if it is voiced or not. The threshold value is estimated as the following procedure:

1. Let $c_3^i(\tau_1, \tau_2)$ be the third-order cumulant of the ith frame, and LC be the logarithm of $c_3^i(\tau_1, \tau_2)$. Here we only compute the special points $c_3^i(0,0) = E\{x(n)^3\}$, and $LC_i = 10\ \log\ 10(c_3^i(0,0))$.

2. Assume the first 10 frames are unvoiced frames. Compute the mean value, $\overline{LC} = (1/10)\sum_{i=1}^{10} LC_i$ and standard deviation, $\sigma_{lc} = ((1/9)\sum_{i=1}^{10}(LC_i - \overline{LC})^2)^{1/2}$ of their LC_i.
3. Compute the threshold value as

$$\text{Threshold} = \overline{LC} + ga \times \sigma_{lc} \quad (4)$$

where ga is a variable, and here it is set to 2.3.

If LC_i of the ith frame is greater than the threshold, it is declared a voiced frame, otherwise is an unvoiced frame. For the unvoiced frame, its sample bispectrum is multiplied by a variable q of small value. If the ith frame is a voiced frame, and its adjacent frames (the $i-1$ and $i+1$) are also identified as voiced, its sample bispectrum is given by a weighted average scheme [23] as

$$B_i(\omega_1, \omega_2) = aB_{i-1}(\omega_1, \omega_2) + bB_i(\omega_1, \omega_2) + aB_{i+1}(\omega_1, \omega_2) \quad (5)$$

where $2a + b = 1$ for $a \le b$, $0 \le a \le 0.33$.

After estimating the bispectrum values, the Least Squares Approach in [25] was used to estimate the magnitude of the Fourier transform of each frame. The phase distortion was barely perceived by the human ear. Hence the phase spectrum of each frame was kept unchanged.

5. Results and discussion

The waveforms of speech processed by Wiener filtering, spectral subtraction and bispectrum approach are given in Fig. 3(b)–(d), respectively. The spectrograms of the signals shown in Fig. 3(a)–(d) are given in Fig. 3(e)–(h), respectively. The evaluated speech signal used here is chosen from one typical male volunteer. It is observed that noise reduction is achieved to some degree for speech processed by all of the three algorithms. Especially for bispectrum approach, the power of noise is reduced dramatically as shown in Fig. 3(h), whereas Fig. 3(f) and (g) are still characterized by certain level of noise. Moreover, frequency components in some higher region are removed in Fig. 3(f) and (g), resulting in severe speech distortion. However, for bispectrum approach, the frequency components are preserved well without removing the essential signal information. So, from Fig. 3, it can be inferred that bispectrum algorithm outperforms Wiener filtering and spectral subtraction for the radar speech quality improvement.

Table 1
Average MOS by enhancement algorithm for radar sensor speech. The radar speech was first corrupted by pink, white and babble noises at an SNR of 0 dB, and then enhanced by these algorithms respectively. These enhanced speech materials were recorded, and played out for rating by 10 listeners using MOS. The final MOS results were obtained by averaging the scores over the 10 listeners for each algorithm.

Algorithm	Pink	White	Babble
Wiener filtering	3.26	3.10	3.43
Spectral subtraction	3.37	3.56	3.30
Bispectrum	4.12	3.89	3.95

The results of the listening test are presented in Table 1. It is found that Wiener filtering procedure received scores, around "3", very close to the spectral subtraction. "3" indicates that the speech signals are still perceptible with slightly annoying. We also observe that the bispectrum algorithm consistently yielded the highest scores in three noise conditions almost near to "4", which indicates that the output signals processed by our bispectrum have the highest speech quality and sound more pleasant for the listeners. A great difference in perceptual quality between speech enhanced by bispectrum and the other two can be observed from listening tests. The most likely explanation for the difference is that Wiener filtering and spectral subtraction inevitably yield more speech distortion when these approaches reduce the noise. And this is especially true for spectral subtraction, because speech enhanced by it suffers serious disturbance from musical noise artifact, which is unpleasant by the human auditory system. Whereas the mechanism of bispectrum used to filter out the noise will not produce the similar negative effect.

The results of LSD for different noise types and algorithms are shown in Table 2. It is noticed that the three algorithms lead to a decrease of LSD values in all noise conditions at different SNR levels. Using Wiener filtering, the worst speech enhancement performance is observed clearly. This is due to that it suppresses noise at the expense of severe target signal distortion. Spectral subtraction exhibits lower LSD values in comparison to Wiener filtering. However, its improvement is not significant. The poor performance may result from musical noise, which is introduced by the

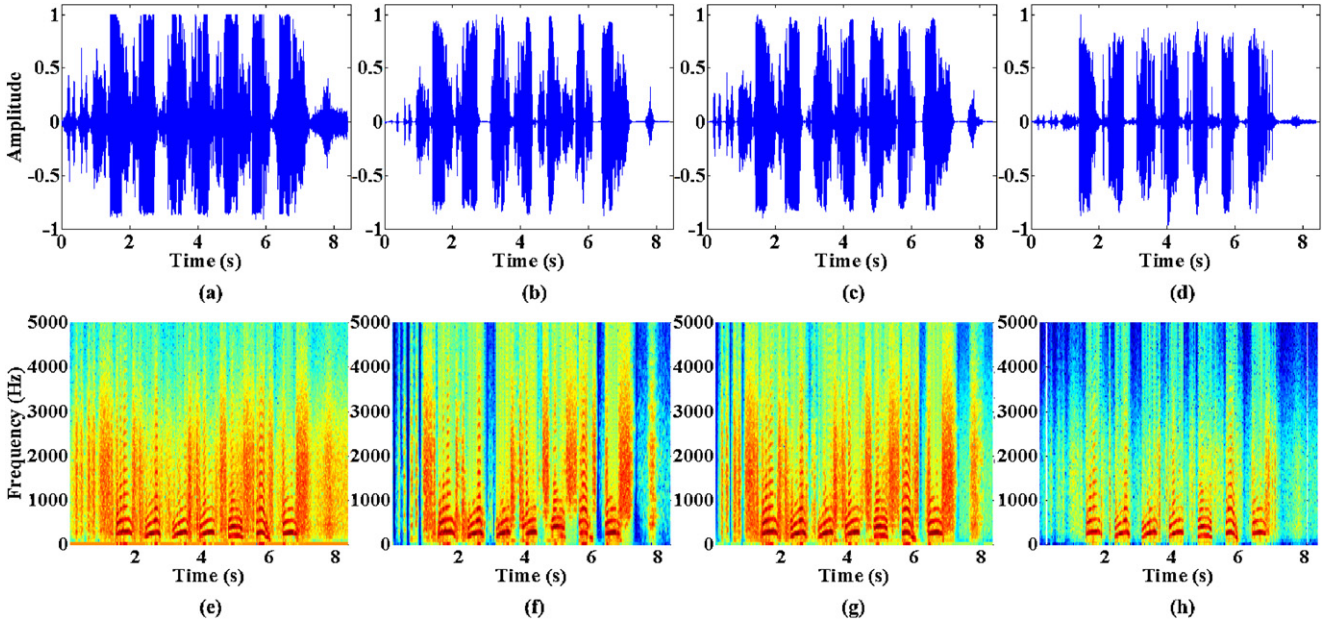

Fig. 3. Results of radar speech enhancement. (a) Original radar speech; (b) speech processed by Wiener filtering; (c) speech processed by spectral subtraction; (d) speech processed by bispectrum; (e)–(h) spectrograms of the signals shown in (a)–(d).

Table 2
LSD of enhanced radar sensor speech signals in the pink, white and babble noise conditions.

Algorithm	Pink				White				Babble			
	0	5	10	20	0	5	10	20	0	5	10	20
Wiener filtering	2.3725	2.6440	2.3808	2.2989	2.1719	2.5775	2.7734	2.4005	1.7522	1.7455	1.8269	1.9993
Spectral subtraction	1.4597	1.4506	1.3461	1.4441	2.1110	1.7932	1.4976	1.4263	1.5782	1.4220	1.3598	1.3722
Bispectrum	1.4568	1.4495	1.1753	0.8664	1.5627	1.4696	1.4738	0.9510	1.5089	1.2803	1.1418	0.9314

algorithm itself. The proposed bispectrum algorithm yields the lowest LSD values for all conditions compared with the other two methods at the same SNR level, which indicates that the speech signal processed by our algorithm has a superior similarity with original speech than those by the other methods. Its remarkable performance can be attributed to the property of bispectrum in that the third-order cumulant of Gaussian noise is zero. Therefore additive Gaussian noise with speech signal can be automatically removed when processed by our algorithm.

In addition, we also considered the effect of gender difference on the performance of our algorithm. The characteristics of vocal tract between male and female differ considerably [26]. This difference has been extensively investigated in the area of acoustics. It is the gender-based difference that has been utilized in the speech related research. For example, Bradlow et al. investigated whether the talker-specific characteristics (e.g., gender, F0 and speaking rate) might affect overall speech intelligibility [27]. Potamitis et al. introduced a speech enhancement technique that explicitly incorporates prior information about the gender or speaker time–frequency characteristics in its formalism [28]. However, different from their study that extracts features carrying gender-based information, our algorithm is based on HOS, which does not introduce individual information correlated with gender. Therefore it is expected that the performance of the enhancement operation is gender independence. To prove this point, we analyzed the LSD for the five males and five females under the same experiment conditions and separated results according to the speaker's gender. We noticed that the results in terms of LSD display no significant difference between male and female. This suggests that gender has no significant impact on the performance of our algorithm.

6. Conclusion

A smart radar sensor was developed to detect the speech signal. This new sensor overcomes the challenges associated with using radar in the measurement of minute organ movement such as small vibrations of the human body around the larynx. The non-contact detection technique also removes the emotional tensions from its user and brings convenient experience for the user. The high operating frequency and superheterodyne receiver combine to improve the measurement sensitivity and directivity. Moreover, bispectrum algorithm was implemented to improve the quality of radar speech. The speech enhancement results show that the algorithm reduces the level of noise and preserves the essential speech information with least distortion. This radar sensor is a promising new device for non-contact speech detection, and also more effort is needed to put into the investigation to further improve the robustness of our system.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) (No. 60571046). The authors also want to thank the participants from the E.N.T. Department, the Xi Jing Hospital for helping with data acquisition and analysis.

References

[1] A. Torkkeli, O. Rusanen, J. Saarilahti, H. Seppä, H. Sipola, J. Hietanen, Capacitive microphone with low-stress polysilicon membrane and high-stress polysilicon backplate, Sensors and Actuators A: Physical 85 (2000) 116–123.
[2] P.R. Scheeper, A.G.H. van der Donk, W. Olthuis, P. Bergveld, A review of silicon microphones, Sensors and Actuators A: Physical 44 (1994) 1–11.
[3] C.H. Pua, S.F. Norizan, S.W. Harun, H. Ahmad, Non-membrane optical microphone based on longitudinal modes competition, Sensors and Actuators A: Physical 168 (2011) 281–285.
[4] M. Pedersen, W. Olthuis, P. Bergveld, A silicon condenser microphone with polyimide diaphragm and backplate, Sensors and Actuators A: Physical 63 (1997) 97–104.
[5] W. Kronast, B. Müller, W. Siedel, A. Stoffel, Single-chip condenser microphone using porous silicon as sacrificial layer for the air gap, Sensors and Actuators A: Physical 87 (2001) 188–193.
[6] S.A. Patil, J.H.L. Hansen, The physiological microphone (PMIC): a competitive alternative for speaker assessment in stress detection and speaker verification, Speech Communication 52 (2010) 327–340.
[7] T. Winkler, S. Pronkine, R. Bardeli, J. Köhler, A study of throat microphone performance in automatic speech recognition on motorcycles, in: NAG/DAGA 2009, Rotterdam, 2009, pp. 1659–1662.
[8] C.M. Santori, S. Jose, Calif, Bone Conduction Microphone – Assembly, U.S. Patent 3,787,641 (1974).
[9] H. Ruser, V. Mágori, Highly sensitive motion detection with a combined microwave-ultrasonic sensor, Sensors and Actuators A: Physical 67 (1998) 125–132.
[10] W. Jianqi, Z. Chongxun, L. Guohua, J. Xijing, A new method for identifying the life parameters via radar, EURASIP Journal on Applied Signal Processing 2007 (2007) 16.
[11] Z.W. Li, Millimeter wave radar for detecting the speech signal applications, International Journal of Infrared and Millimeter Waves 17 (1996) 2175–2183.
[12] J. Holzrichter, G. Burnett, L. Ng, W. Lea, Speech articulator measurements using low power EM-wave sensors, Journal of the Acoustical Society of America 103 (1998) 622–625.
[13] R. Hu, B. Raj, A robust voice activity detector using an acoustic Doppler radar, in: 2005 IEEE Workshop on Automatic Speech Recognition and Understanding, 2005, pp. 319–324.
[14] S. Boll, Suppression of acoustic noise in speech using spectral subtraction, IEEE Transactions on Acoustics, Speech and Signal Processing 27 (1979) 113–120.
[15] M. Berouti, R. Schwartz, J. Makhoul, Enhancement of speech corrupted by acoustic noise, in: ICASSP 79 IEEE International Conference on Acoustics Speech and Signal Processing, vol. 4, 1979, pp. 208–211.
[16] B. Moran, Mathematics of radar, in: J.S. Byrnes (Ed.), Twentieth Century Harmonic Analysis – A Celebration, NATO Sci. Ser. II Math. Phys. Chem., Kluwer Acad. Publ., Dordrecht, 2001, pp. 295–328.
[17] S. Han, S. Jang, Y. Woo, J. Lee, Third-order moment spectrum and weighted fuzzy classifier for robust 2-D object recognition, Computers & Mathematics with Applications 45 (2003) 699–707.
[18] M. Jiao, G. Lu, X. Jing, S. Li, Y. Li, J. Wang, A novel radar sensor for the non-contact detection of speech signals, Sensors 10 (2010) 4622–4633.
[19] B. Jounghoon, R.H. Baran, K. Hanseok, Dual channel based speech enhancement using novelty filter for robust speech recognition in automobile environment, IEEE Transactions on Consumer Electronics 52 (2006) 583–589.
[20] S. Ahmadi, A.S. Spanias, Low bit-rate speech coding based on an improved sinusoidal model, Speech Communication 34 (2001) 369–390.
[21] C.L. Nikias, J.M. Mendel, Signal processing with higher-order spectra, IEEE Signal Processing Magazine 10 (1993) 10–37.
[22] S.A. Dianat, R.M. Rao, Fast algorithms for phase and magnitude reconstruction from bispectra, Optical Engineering 29 (1990) 504–512.
[23] R. Fulchiero, A.S. Spanias, Speech enhancement using the bispectrum, in: ICASSP93 IEEE International Conference on Acoustics Speech and Signal Processing, vol. 484, 1993, pp. 488–491.
[24] G. Sundaramoorthy, Improved Techniques for Bispectral Reconstruction of Signals, M.D. thesis, Rochester, New York, August 1990.
[25] G. Sundaramoorthy, M.R. Raghuveer, S.A. Dianat, Bispectral reconstruction of signals in noise: amplitude reconstruction issues, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (1990) 1297–1306.
[26] Y. Lu, M. Cooke, The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise, Speech Communication 51 (2009) 1253–1262.
[27] A.R. Bradlow, G.M. Torretta, D.B. Pisoni, Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics, Speech Communication 20 (1996) 255–272.

[28] I. Potamitis, N. Fakotakis, G. Kokkinakis, Gender-dependent and speaker-dependent speech enhancement, in: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2002, pp. I-249–I-252.

Biographies

Ying Tian received her M.E. degree in communication engineering from Engineering University of Armed Police Force, Xi'an, China, in 2011. She is currently a Ph.D. student in biomedical engineering at The Fourth Military Medical University. Her research interest is biomedical signal noncontact detecting and processing.

Sheng Li was born in Sinkiang, China, on January 26, 1972. He received the Ph.D. degree in biomedical engineering at Xi'an Jiaotong University of China, Xi'an, in 2005. His research interests include radar wave biomedical signal noncontact detecting and processing, speech production and enhancement.

Hao Lv received his Ph.D. degree in biomedical engineering from The Fourth Military Medical University (FMMU), Xi'an, China, in 2010. He is currently teaching in the Faculty of Biomedical Engineering, FMMU. His research interest is biomedical signal noncontact detecting and processing.

Jianqi Wang was born in April, in 1962, China. He received the diploma in information and control engineering from Xi'an Jiaotong University, China, in 1984. From 2002 to 2006, he was the doctor student of the Key Laboratory of Education Ministry of China, Xi'an Jiaotong University. From 1996 to 2006, he was teaching in the Faculty of Biomedical Engineering, FMMU. He is currently serving as Professor and Vice Director of the Faculty of Biomedical Engineering at the same university. His research interest is biomedical signal noncontact detecting and processing.

Xijing Jing was born in 1957 in China. He received the diploma in electrical engineering from Xi'an Jiaotong University, China, in 1978. From 1987 to 2006, he was teaching in the Faculty of Biomedical Engineering of FMMU. He is currently serving as an Associate Professor at the same university. His research interest is military medical engineering.