

Remote audio signals detection using a partial-fiber laser Doppler vibrometer

Shi-song Wu^{a,b}, Tao Lv^{a,b}, Xi-yu Han^{a,b}, Chun-hui Yan^{a,b}, He-yong Zhang^{a,*}

^a Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, State Key Laboratory of Laser Interaction with Matter, Changchun 130033, China

^b University of the Chinese Academy of Sciences, Beijing 10039, China

ARTICLE INFO

Keywords:

LDV
Laser hearing
Partial-fiber

ABSTRACT

All-fiber laser Doppler vibrometer (LDV) systems have great potential in the application of remote acoustic detection. However, the system performance is impaired due to the serious intermediate frequency (IF) crosstalk signal, which is caused by the fiber circulator. In this paper, a system adopting a partial-fiber structure is proposed and analyzed. Simulative and experimental analysis indicate that the IF crosstalk signal is mainly caused by the fiber circulator, and it damages demodulation signal. In order to reduce the IF crosstalk signal, a polarization prism is employed to substitute the circulator. Experimental results reveal that the IF crosstalk signal can be eliminated by the partial-fiber LDV effectively. The comprehensible speech signals within the range of 50m can be obtained.

1. Introduction

LDV has the characteristics of long distance, non-contact and high sensitivity. It has been widely used in industry and military field [1,2]. On the one hand, LDV can detect and measure extremely tiny vibration of a target at a long distance; on the other hand, the acoustic pressure can vibrate the objects near the audio sources. Therefore, the voice signals of a human could be acquired by capturing the vibration of a target's surface. Li and Wang [3–7] have presented their researches in detecting and processing voice signals of people from large distances using a LDV from Polytec. However, the sensor heads are usually bulky and heavy because of the inner separated structure of the commercial LDV systems (e.g., the Polytec OFV 505 system has dimensions of 120 mm × 80 mm × 345 mm and weight of 3.4 kg). An all-fiber LDV system has the advantages in smaller size, lightweight design, and more robust structure, therefore it is less prone to structural vibrations and more suitable for remote speech detection [8]. In previous study [9,10], we have established an all-fiber LDV to detect remote voice signals. Although this system can acquire more than 120 m comprehensible voice signals, the serious IF crosstalk signal impairs system performance. In this paper, a partial-fiber LDV is developed by employing a polarization prism to eliminate the IF crosstalk signal. A series of experiments are carried out to verify the capability of partial-fiber LDV system. Experimental results indicate that the IF crosstalk signal can be eliminated by the partial-fiber LDV efficiently. Besides, the

comprehensible speech signals within the range of 50 m can be obtained.

2. Experimental setup

The schematic diagram of a traditional LDV is illustrated in Fig. 1. The LDV is composed of optical unit and electrical unit. A 20-mW single mode CW laser with the line-width less than 10 kHz at wavelength of 1550 nm is adopted as a transmitter. The output laser beam is divided into two beams by an optical fiber splitter. One beam is taken as the local-oscillator (LO) beam after passing through the acousto-optic frequency shifter (AOFS), which shifts the frequency of the LO beam by ω_{AO} . Meanwhile, the other beam is focused on the detected target after passing through a fiber circulator and a telescope. The Doppler frequency shift of the received beam is generated due to the vibration of the target, which can be expressed as

$$\varphi(t) = 4\pi \frac{S(t)}{\lambda} \quad (1)$$

where $S(t)$ is the vibration displacement, λ is the wavelength. Being received through the telescope and coupled into the fiber circulator, the reflected beam is mixed with the LO beam in a polarization-maintaining fiber coupler to produce a beat signal, which is converted into a voltage signal by a photoelectric balanced detector. The IF signal is emerged when the voltage signal passes a bandpass filter. The LO signal, echo

* Corresponding author.

E-mail address: zhanghy@ciomp.ac.cn (H.-y. Zhang).

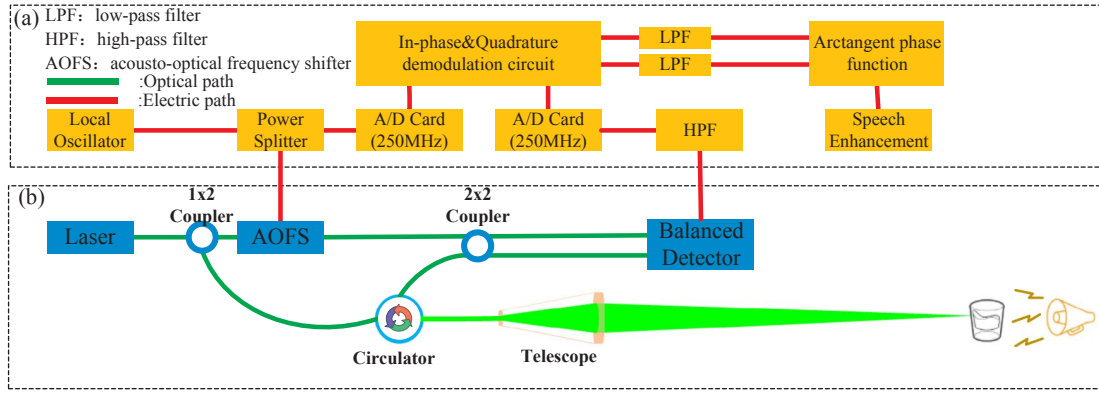


Fig. 1. Schematic diagram of the LDV (a) electrical unit (b) optical unit.

signal and IF signal are expressed as follows:

$$I_{LO} = A_{LO} \cos[(\omega_c + \omega_{AO})t + \varphi_1] I_S = A_S \cos[\omega_c t + \varphi(t) + \varphi_2] u_{IF} \\ = \alpha A_{LO} A_S \cos[\omega_{AO} t + \varphi(t) + \varphi_1 - \varphi_2] \quad (2)$$

where A_{LO} , A_S are the amplitude of local-oscillator and signal beam respectively, ω_{AO} is the frequency shift caused by AOFS, φ_1 and φ_2 are random phases, α is the photoelectric conversion efficiency, $\varphi(t)$ is the Doppler shift.

In order to obtain the characteristic of the vibrating target, it's a classical method to demodulate the beat signal using quadrature demodulation and arctangent phase algorithm [11]. This method has been analyzed and tested in recent years, enabling the realization of high-accuracy heterodyne signal processing in commercial LDV.

The demodulation and arctangent phase algorithm block diagram is depicted as Fig. 2.

The IF signal is divided into two equal parts by the power splitter. One is mixed with half of the LO beam, transformed into the baseband signal u_I after passing through a low-pass filter. While the other is mixed with a 90° phase shift of the rest LO beam and changed into the baseband signal u_Q after passing through a same low-pass filter. The expression of the baseband signal is given as follows:

$$u_I \approx \alpha A_{LO} A_S \cos[\varphi(t) + \varphi_1 - \varphi_2] u_Q \approx \alpha A_{LO} A_S \sin[\varphi(t) + \varphi_1 - \varphi_2] \quad (3)$$

It is simple to get the Doppler frequency shift $\varphi(t)$ using arctangent phase algorithm. Once the Doppler frequency shift $\varphi(t)$ is calculated, the speech signal can be reconstructed. The ambiguity of the arctangent function can be removed by phase unwrapping algorithm. The Doppler frequency shift $\varphi(t)$ can be simplified as:

$$\varphi(t) = \arctan(u_Q/u_I) + m\pi + \Delta\varphi \quad (4)$$

There is a channel crosstalk in the fiber circulator due to structural defects of the fiber circulator in the above typical diagram of LDV. Because of the existence of channel crosstalk, the echo beam and LO beam change their expressions:

$$I_{LO} = A_{LO} \cos[(\omega_c + \omega_{AO})t + \varphi_1] I_S = A_S \cos[\omega_c t + \varphi(t) + \varphi_2] \\ + A_c \cos[\omega_c t + \varphi_3] \quad (5)$$

Therefore, baseband signal u_I and u_Q become the following expressions:

$$u_I \approx \alpha A_{LO} A_S \cos[\varphi(t) + \varphi_1 - \varphi_2 - \varphi_3] + \alpha A_{LO} A_c \cos[\varphi_1 - \varphi_2 - \varphi_3] u_Q \\ \approx \alpha A_{LO} A_S \sin[\varphi(t) + \varphi_1 - \varphi_2 - \varphi_3] + \alpha A_{LO} A_c \sin[\varphi_1 - \varphi_2 - \varphi_3] \quad (6)$$

Assuming the phase difference $\varphi_1 - \varphi_2 - \varphi_3$ is equal to 0, the Doppler frequency shift is obtained by adopting arctangent phase algorithm.

$$\varphi(t) = \arctan\left(\frac{u_Q}{u_I}\right) = \arctan\left(\frac{\sin(\varphi(t))}{\cos(\varphi(t)) + \frac{A_c}{A_S}}\right) \\ = \arctan\left(\tan(\varphi(t)) \left(1 + (-1)^n \sum_{n=1}^{\infty} \left(\frac{A_c}{A_S}\right)^n \left(\frac{1}{\cos(\varphi(t))}\right)^n\right)\right) \quad (7)$$

The crosstalk of the fiber circulator is negligible when the intensity of crosstalk signal is much less than the one of echo signal, ie $A_c \ll A_S$. However, when the two are comparable in magnitude or the intensity of crosstalk is greater than the one of echo, the crosstalk of the fiber circulator affects the measurement accuracy of Doppler frequency shift directly.

3. Simulation

According to laser range equation, the power of echo signal P_S is derived as follows:

$$P_S = \frac{P_T}{4\pi R^2} \times \frac{\sigma}{4\pi R^2} \times \frac{\pi D^2}{4} \times \eta_{atm} \eta_{sys} \quad (8)$$

where P_T is the power of transmitted laser, σ is the radar cross section of target, D is the receiving aperture, R is the distance between laser transmitter and target, η_{atm} is atmospheric transmission coefficient, η_{sys} is transmission coefficient of the optical system. Since the laser is focused on the target through focusing system, all of the radiant energy near the spot can be reflected. Therefore, the target can be taken as an expansion target. The area of the irradiation is expressed as:

$$dA = \frac{\pi R^2 \theta_T^2}{4} \quad (9)$$

where θ_T is the diffraction limit angle of the laser beam. For extended Lambert scattering targets, there are:

$$\sigma_{ext} = 4\rho_{ext} R^2 \theta_T^2 \quad (10)$$

The power of echo signal P_S can be simplified as:

$$P_S = \frac{\pi P_T \rho_{ext} D^2}{(4R)^2} \eta_{atm} \eta_{sys} \quad (11)$$

where ρ_{ext} is the average reflective coefficient of target.

Assuming the measured target is file folder, the received echo beam power P_S equals 0.318 nW theoretically with the condition of $P_T = 18$ mW, $R = 30$ m, $D = 1.5$ cm, $\eta_{atm} = 0.9$, $\eta_{sys} = 0.8$, $\rho_{ext} \approx 0.5$. In the system, the channel crosstalk level of the fiber circulator is about 60 dB and the power of the crosstalk equals 18 nW. Supposing that movement of target is a simple harmonic vibration ($S(t) = 50 \times 10^{-6} \times \cos(2\pi \times 500 t)$), simulation is carried out based on the following parameters in Table 1.

Compared Fig. 3 with Fig. 4, it is obvious that the IF crosstalk can directly affect the accuracy of measurement when the intensity of

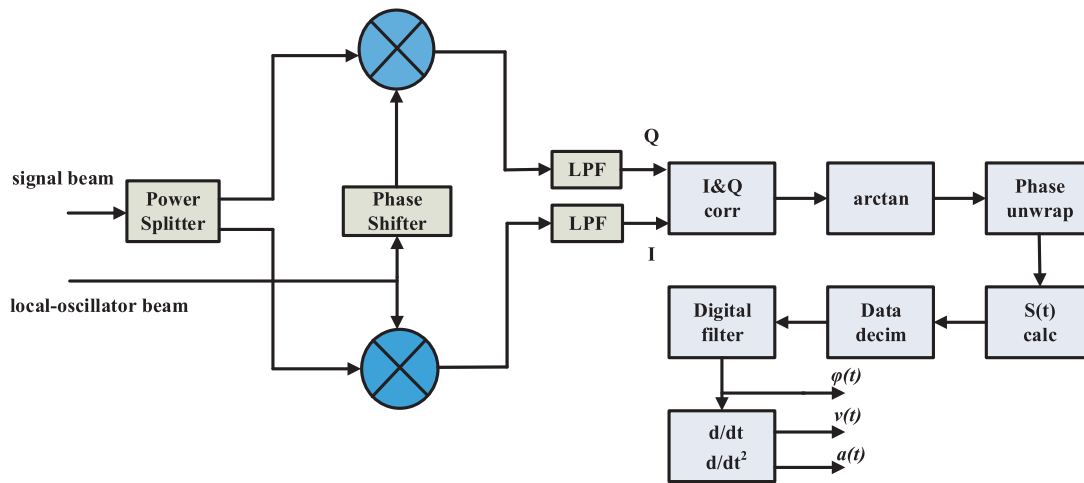


Fig. 2. The block diagram of demodulation and arctangent phase algorithm.

Table 1
Parameters of the simulation.

Parameters	Value
Wavelength	1550 nm
Power	20 mW
Linewidth	10 kHz
A_{LO}	2507
A_S	1
A_c	7.5
$\Delta\varphi$	0
α	1
Sampling rate	320 MHz
Effective range	30 m

crosstalk signal is much larger than the one of echo signal. In fact, the echo intensity is so weaker than one of the crosstalk signal that the impact of IF crosstalk on the system cannot be ignored because of the far detection range as well as low reflectivity of the target. Although it is possible to reduce the crosstalk by band pass filtering, adaptive trap filtering, etc., it is bound to cause unpredictable damage to the useful components of the signal.

4. Experimental improvement

In order to eliminate the IF crosstalk in the all-fiber LDV system, a polarization prism is employed to substitute the circulator (see Fig. 5).

To verify the capability difference between all-fiber system and partial-fiber system, an experiment is set up. As depicted in Fig. 6a, a

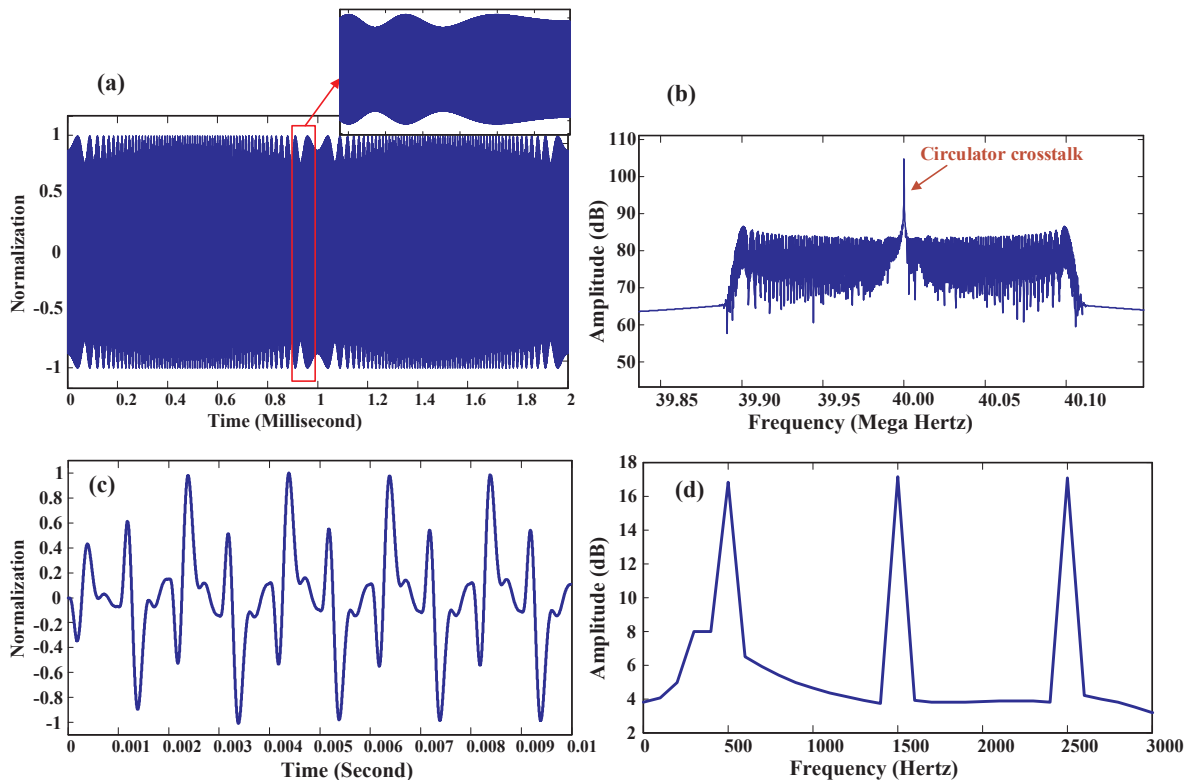


Fig. 3. Simulation results when crosstalk exits (a) IF signal u_{IF} in time domain (b) IF signal u_{IF} in frequency domain (c) Doppler frequency shift $\varphi(t)$ in time domain (d) Doppler frequency shift $\varphi(t)$ in frequency domain.

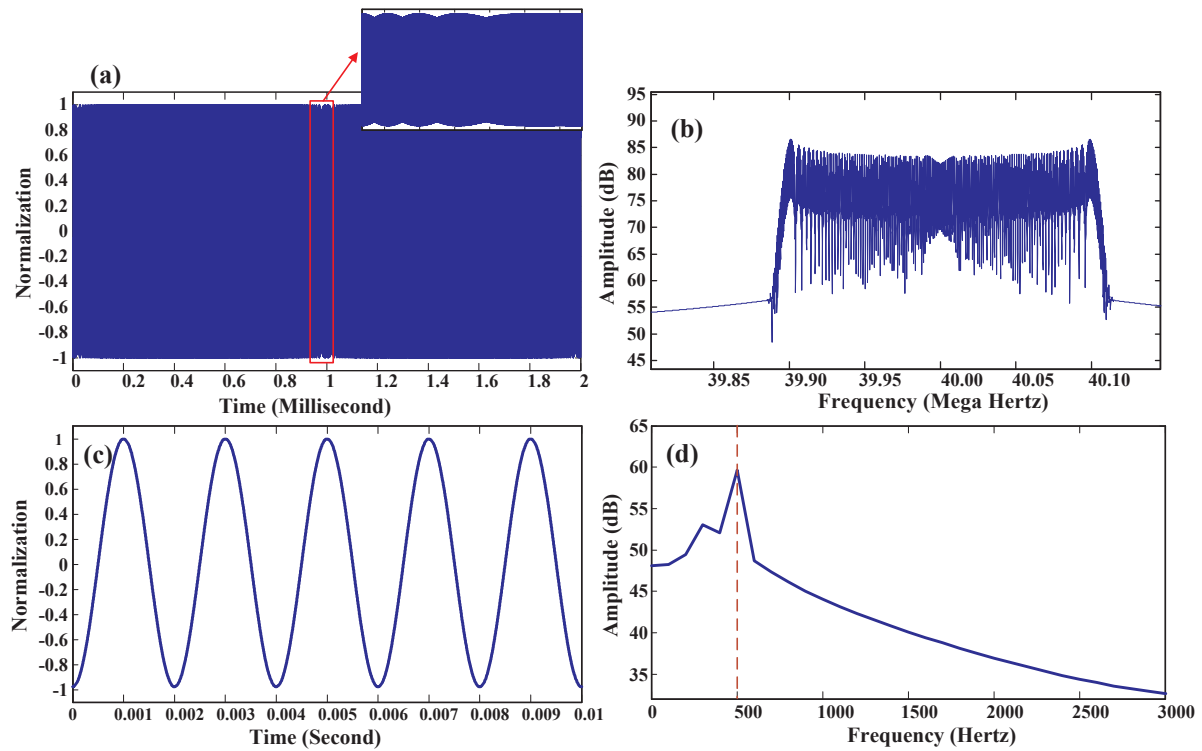


Fig. 4. Simulation results when crosstalk does not exit (a) IF signal u_{IF} in time domain (b) IF signal u_{IF} in frequency domain (c) Doppler frequency shift $\varphi(t)$ in time domain (d) Doppler frequency shift $\varphi(t)$ in frequency domain.

file folder is taken as the target. It is forced to vibrate under the effect of loudspeaker enclosure, which is derived by single-frequency sinusoidal tone generated by signal generator. The frequency of the single tone is adjustable by toning the signal generator. The LDV transmitted the laser beam perpendicularly to the file folder surface so as to gain the optimum reflected signal.

The experimental results are demonstrated in Figs. 7 and 8. It is obvious that there is a relatively strong IF crosstalk signal which caused by the circulator in the all-fiber system (see Fig. 8a). This crosstalk interferes the demodulation results significantly (see Fig. 8b and c). However, the partial-fiber LDV has ability to eliminate the IF crosstalk signal effectively (see Fig. 7a), and the single tone can be acquired and calculated (see Fig. 7b and c).

To further indicate the partial-fiber LDV's acquiring capability of remote speech signals, the experiment is set up as Fig. 6b. A speaker is located at a distance of about 50 m from the LDV and 30 cm from the target (a mineral water bottle). The mineral water bottle is forced to vibrate due to the speaker's voice and its frequency equals the sound-

field frequency. The voice can be acquired by detecting the vibration of the mineral water bottle.

To evaluate the experiment result, both subjective and objective evaluations are implemented. The subjective evaluation is named Mean Opinion Score (MOS) evaluation criterion (see Table 2), and the objective evaluation is named spectrogram/waveform comparison. Fig. 9 shows the spectrogram and waveform of partial-fiber LDV speech signals (see Fig. 9b), and corresponding clean signals (see Fig. 9a) captured at the same time by a voice recorder. From the spectrogram comparison, we can find the partial-fiber LDV audio signal is degraded by a strong noise, because the inherent "speckle pattern" problem on a normal "rough" surface as well as circuit noise creates an undesirable loud noise. The valid frequency involved in the clean signal is same with the LDV audio signal. Moreover, the waveform of LDV signal is similar to the clean signal. Besides ten volunteers are invited to evaluate the performance of the system under the MOS evaluation criterion. The average of the evaluation result is 2.1. Although the MOS of the signal is not very high, the meaning of voice is still understandable. Experiments

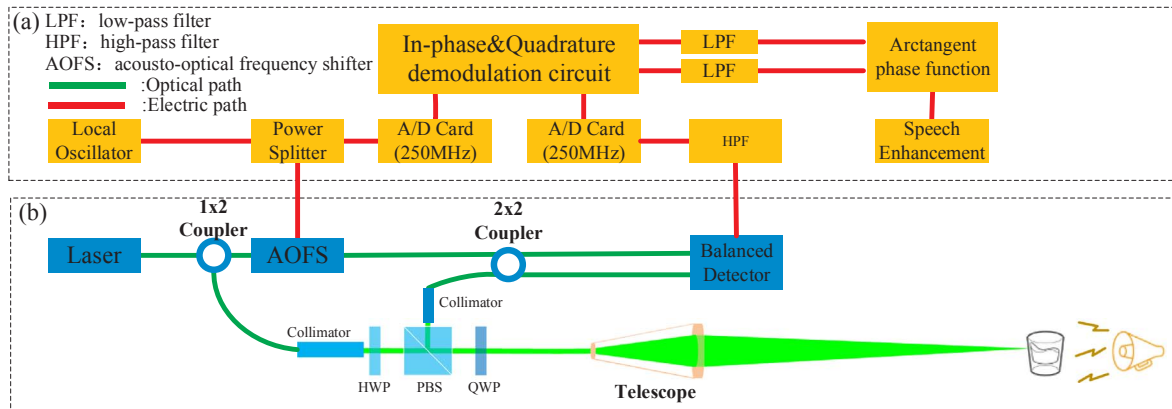


Fig. 5. Improvement on schematic diagram of the LDV (a) electrical unit (b) optical unit.

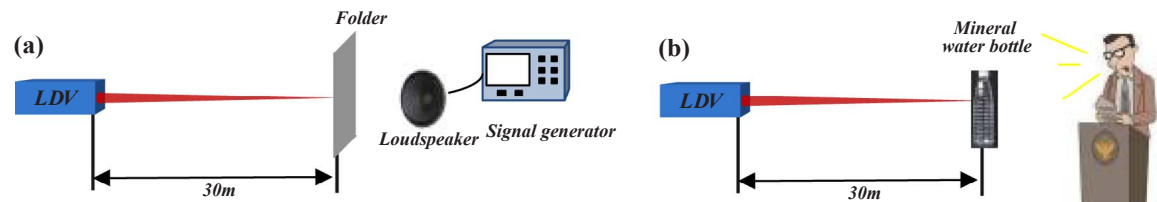


Fig. 6. (a) Experimental block diagram of detecting single tone (b) Experimental block diagram of detecting voice.

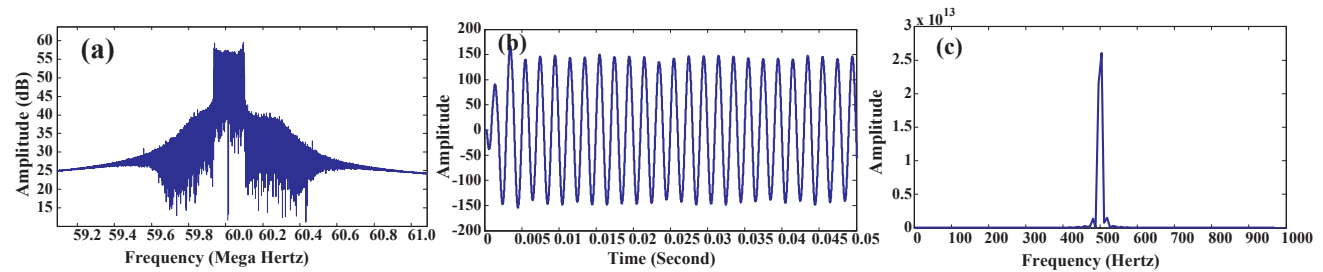


Fig. 7. Experimental results of the partial-fiber system (a) spectrum of the IF signal (b) waveform of the Doppler frequency shift (c) spectrum of the Doppler frequency shift.

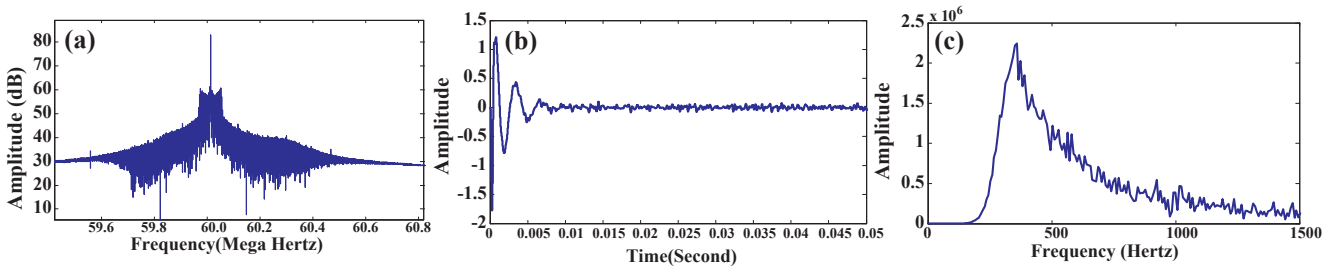


Fig. 8. Experimental results of the all-fiber system (a) spectrum of the IF signal (b) waveform of the Doppler frequency shift (c) spectrum of the Doppler frequency shift.

Table 2
Mean opinion score evaluation criterion.

The score of MOS	Intelligibility	Distortion level
5	Excellent	Don't feel distortion
4	Good	Just feel distortion
3	Okay	Feel distortion and have a little disgust
2	Difficult	Obviously feel distortion but can bear it
1	Bad	Unacceptable

indicate that the comprehensible speech signal within the range of 50 m can be obtained by the partial-fiber LDV.

5. Conclusion

All-fiber LDV system has great potential in remote acoustic detection applications. However, the performance of the LDV system is degraded because of disturbance of the IF crosstalk signals. In this paper,

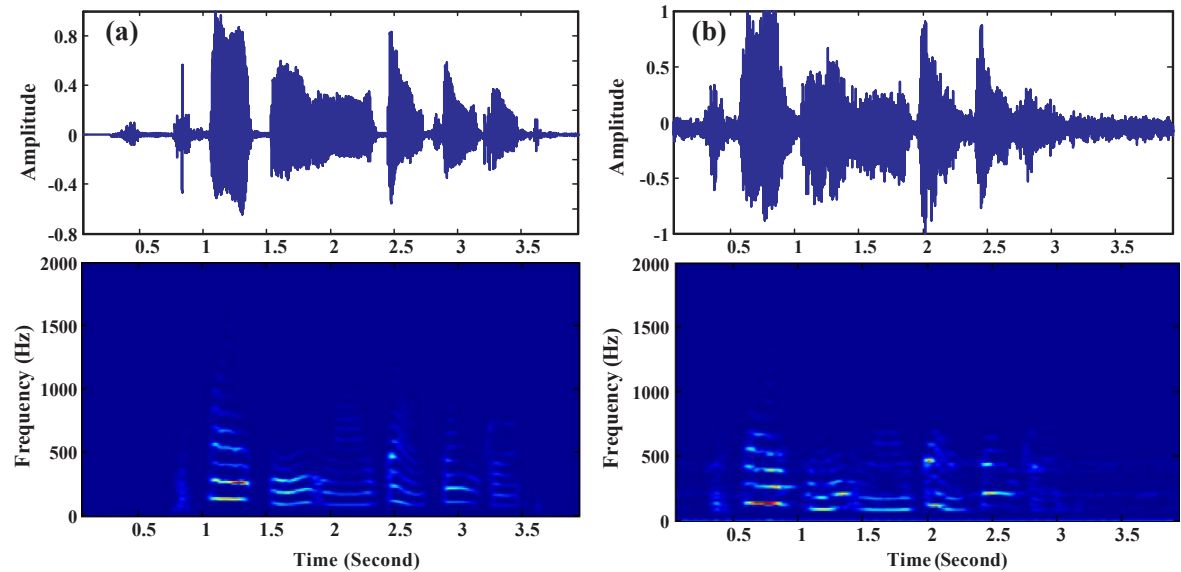


Fig. 9. (a) Waveform and spectrogram of clean voice (b) waveform and spectrogram recorded by partial-fiber LDV system.

the analytical expression of the Doppler frequency shift affected by the crosstalk of the fiber circulator is derived. Simulation is carried out to confirm the effect of crosstalk on demodulation. A partial-fiber LDV is developed to eliminate the IF crosstalk signal. Experimental results reveal that the IF crosstalk signal can be eliminated by the partial-fiber LDV effectively so that the comprehensible speech signals within the range of 50 m can be obtained easily. This system can be applied to various applications such as rescue in disaster and remote area surveillance.

Acknowledgement

This work is supported by the National Natural Science Foundation of China under Grant No. 61205143.

References

- [1] Shang Jianhua, He Yan, Liu Dan, Zang Huaguo, Chen Weibiao. Laser Doppler vibrometer for real-time speech-signal acquirement. *Chin Opt Lett* 2009;7(8):080732.
- [2] Bauer Martin, Ritter Frank, Siegmund Georg. High-precision laser vibrometers based on digital Doppler-signal processing. *Proc of SPIE* 2002;4827:50–61.
- [3] Qu Y, Wang T, Zhu Z. Remote audio/video acquisition for human signature detection. *IEEE Comput Soc Conf Comput Vision Pattern Recognition Workshops* 2012;2009:66–71.
- [4] Li W, Liu M, Zhu Z, et al. LDV remote voice acquisition and enhancement. *Int Conf Pattern Recognit* 2006:262–5.
- [5] Wang AT, Zhu Z, Divakaran A. Long range audio and audio-visual event detection using a laser doppler vibrometer. *Evolutionary Bio-Inspired Comput: Theory Appl IV* 2010;7704. 77040J-77040J-6.
- [6] Qu Y, Wang T, Zhu Z. Remote audio/video acquisition for human signature detection. *Cvpr'09 Biometrics* 2009:66–71.
- [7] Qu Y, Wang T, Zhu Z. An active multimodal sensing platform for remote voice detection. In: *Advanced intelligent mechatronics (AIM)*, 2010 IEEE/ASME international conference on; 2010. p. 627–32.
- [8] Li Rui, Madampoulos Nicholas, Zhu Zhigang, Xie Liangping. Performance comparison of an all-fiber-based laser Doppler vibrometer for remote acoustical signal detection using short and long coherence length lasers. *Appl Opt* 2012;51(21):5011–8.
- [9] Zhang He-yong, Lv Tao, et al. A novel role of arctangent phase algorithm and voice enhancement techniques in laser hearing. *Appl Acoust* 2017;126:136–42.
- [10] Lv Tao, Zhang He-yong, Yan Chun-hui. Double mode surveillance system based on remote audio/video signals acquisition. *Appl Acoust* 2018;129:316–21.
- [11] Bauer M, Ritter F, Siegmund G. High-precision laser vibrometers based on digital Doppler signal processing. In: *Fifth international conference on vibration measurements by laser techniques*; 2002. p. 50–61.