Optics Communications 331 (2014) 6–13

Contents lists available at ScienceDirect

Optics Communications

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

Audio signal reconstruction based on adaptively selected seed points from laser speckle images

Ziyi Chen [a], Cheng Wang [a,*], Chaohong Huang [a], Hongyan Fu [a], Huan Luo [a], Hanyun Wang [b]

[a] Centre of Excellence for Remote Sensing and Spatial Informatics, School of Information Science and Engineering, Xiamen University, Xiamen, Fujian 361005, China
[b] School of Electron Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China

ARTICLE INFO

Article history:
Received 29 March 2014
Received in revised form
13 May 2014
Accepted 15 May 2014
Available online 28 May 2014

Keywords:
Laser speckle pattern
Vibration recovery
Audio signal reconstruction
High-speed camera

ABSTRACT

Speckle patterns, present in the laser reflection from an object, reflect the micro-structure of the object where the laser is illuminated on. When the object surface vibrates, the speckle patterns move accordingly, and this movement can be recovered with a high-speed camera system. Due to the low signal to noise ratio (SNR), it is a challenging task to recover the micro-vibration information and reconstruct the audio signal from the captured speckle image sequences fast and effectively. In this paper, we propose a novel method based on pixels' gray value variations in laser speckle patterns to work out with the challenging task. The major contribution of the proposed method relies on using the intensity variations of the adaptively selected seed points to recover the vibration information and the audio signal with a novel model that effectively fuses the multiple seed points' information together. Experiments show that our method not only recovers the vibration information with high quality but is also robust and runs fast. The SNR of the experimental results reach about 20 dB and 10 dB at the detection distances of 10 m and 50 m, respectively.

1. Introduction

Nowadays, laser speckle pattern has many application scenarios, such as information extraction in medical field [1–3], tilt and displacement measurements [4,5], vibration simulation [6], etc. Since sound detection via optical ways is simple to setup, it is becoming increasingly attractive to detect sound via laser speckle patterns. It is of great significance to detect sound from a remote position in security domain, military application, and criminal monitoring, etc. Detecting sound with laser speckle patterns is easier to achieve under some situations than using electronic detectaphone, as it does not need to put any detection devices nearby the targets. It is able to retrieve the micro-vibration on the object's surface through speckle interferometry, but the speckle interferometry method suffers several disadvantages. First, the detectable distance is limited. Second, there is a limit about the minimal detectable displacements of the speckles. Third, the interference needs the reference light and the condition is rigorous. At first, sounds were normally detected by collecting and detecting the light reflected from the window using an optical interferometer [7,8]. These sound detection methods suffer from the following four major disadvantages [9]: first, as all sounds are detected together, the methods must apply digital blind source separation post-processing algorithms. Second, the projection laser and the detection interferometer module must be placed in very specific positions such that, indeed, the reflected beam is directed towards the detection module. Third, the detection module is complicated and sensitive to errors as they are all interferometer based configurations. Fourth, the methods require a window to be positioned near the voice source.

To overcome the above disadvantages, Zalevsky et al. [9] proposed vibration information extraction through displacements between adjacent speckle patterns (VIETD). In VIETD, a speckle pattern is first generated by projecting a laser beam on the surface of an object, and then the speckle pattern's movements are recorded via a defocused high-speed camera system. After that, the audio signal is reconstructed according to the movements that reflect the vibration information of the laser-pointed object. Furthermore, the quality of the restructured sound signal can be improved by denoising methods [10]. However, the VIETD method has several limitations in applications. First, the image matching procedure is time-consuming. Second, the accuracy of image matching depends on the quality of the obtained speckle images. Since the image quality is generally declined with increasing distance between the image sensor and the object, and the requirement of clear structure information of the speckle images leads to the reduction of the received reflection power, system's detectable distance may be limited. Third, the displacement

---

* Corresponding author.
E-mail address: cwang@xmu.edu.cn (C. Wang).

http://dx.doi.org/10.1016/j.optcom.2014.05.038
0030-4018/© 2014 Elsevier B.V. All rights reserved.

between the sequential speckle images is usually within only several pixels, thus the estimation accuracy of the vibration may not be accurate enough.

Based on similar speckle pattern theory, optical devices [11–13] with single photo-detectors were developed to extract the vibration information of object's surface in real-time. These devices are able to detect the oscillations of either whole speckle patterns [11,12] or single speckles [13]. Veber et al. recently proposed a method that directly obtains sound information from analog signals generated by a masked single photo-detector system [14]. The major limitation of Veber's work is that the pattern of the mask needs to be well designed, fabricated and adjusted according to the shapes and the sizes of the speckle patterns, which increases the difficulty for using. In fact, it is difficult to adjust the mask shape to the speckle pattern as the shape of the speckle pattern is influenced by the following facts: the imaging distance, the micro-structure of the illuminated object surface, the wavelength of the laser beam, and the focal length of the observe lens. In the view of energy, speckle pattern's energy decreases when the speckles pass through the mask or the spatial filter, leading to reducing the detectable distance.

In this paper, we propose a novel micro-vibration detection method based on the laser speckle images recorded by a high-speed camera, which needs neither a mask nor a strict defocus level constraint. There are three steps in our proposed method. First, we capture the speckle image sequences via a high-speed camera. Second, we adaptively select some special pixels in the speckle images as seed points according to some rules that will be analyzed in Section 2.2. Third, we compute the gray-value variation information of the seed points, and finally fuse the variation information with a novel model for audio signal reconstruction.

Due to utilization of gray value variations from the adaptively selected pixels instead of the displacement between adjacent images to reconstruct the vibration, the processing of the proposed method is faster than VIETD since there is no need to do heavy-computation displacement estimation. And the proposed method is more accurate and sensitive as the dynamic gray value variation range of a pixel is up to 100 times larger than the displacements in the high-speed image sequences. Besides, the proposed method does not need a mask, thus the sensors can obtain higher SNR speckle pattern signals. In the experiments, we also show that the proposed method has flexible constraints to the lens defocus level, which are a key factor for obtaining good quality of speckle images under far detection distance. The last key contribution in the paper is the proposed novel fusion model that fuses the multiple seed points' information together to guarantee and improve the performance of the proposed method.

Section 2 describes the theoretical model and the computational complexity of the proposed method, Section 3 shows the experimental results and discussions, and Section 4 draws the concluding remarks. Finally, the acknowledgment and the references are given.

2. Theoretical model

2.1. Experimental setup and system framework

The schematic diagram of the proposed system's framework is shown in Fig. 1, which includes a laser projector (MSL-III-532-50 mW), a lens (a Nikon lens whose focal length can vary ranging from 30 mm to 300 mm), and a high speed camera (Pixelink-74L148). The laser projector emits a laser beam onto the surface of the object of interest (OOI), generating the speckle field in the space. The high speed camera records the speckle patterns via the lens at a high frame rate, e.g. 2–8K fps, and the vibration information of the OOI surface is recovered from the recorded speckle image sequence (the sequence length is equal to frame rate multiply detection time) via our post processing method. In Fig. 1, L is the distance between speckle pattern plane being observed and the surface of OOI, U is the distance between the plane of speckle patterns and the camera lens, f is the focus length, D is the diameter of laser spot, and α is the tilt angle of OOI caused by micro-vibrations.

Fig. 2 shows the flow chart of the proposed method. In the first step, we project the laser on the surface of OOI. In the second step, the high-speed camera records the speckle images with high frame rate. After recording the speckle image sequences, we adaptively select some points as seed points (discussed in Sections 2.2 and 2.3), and finally the vibration information and the audio signal are reconstructed through fusing the gray value variations of multiple seed points (discussed in Section 2.4).

2.2. Principle analysis and seed-selection criteria

The rigid movements of OOI's surface include tilt, transverse and axial translations. Generally, the three types of movements are difficult to separate. However, according to Ref. [9], for the far-field speckle, the transverse and axial translations of OOI's surface have barely detectable effect on speckle image under the condition of strongly defocusing, thus the tilt movement can be separated from the translation movements. As a result, the imaged speckle pattern shows only a displacement that originates from the tilt movement and can be calculated in terms of the tilt angle and the defocusing level. Ascribe to the movement separation, the vibration of OOI's surface can be reconstruction directly from the displacements of recorded high-speed speckle sequences as done by Zalevsky et al. [9]. One key problem of Zalevsky's method is that the estimation algorithm of speckle displacements, such as a cross-correlation method, is generally of low efficiency, especially for sub-pixel

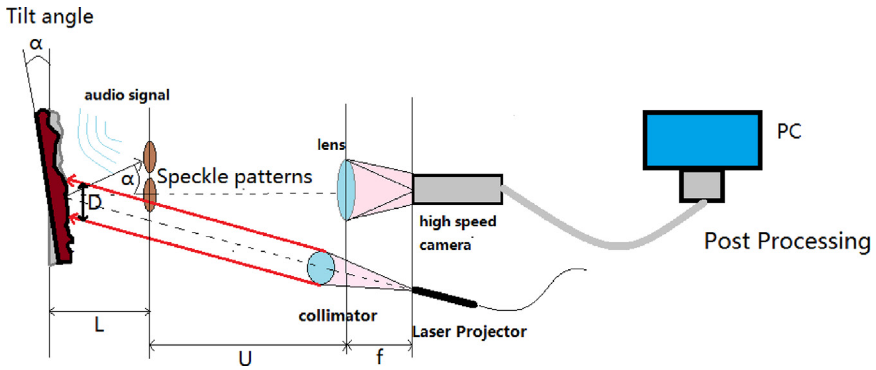

Fig. 1. An overview of the proposed system framework.