

DOI:10.3969/j.issn.1003-5060.2023.02.007

无人机晃动下基于相位的视觉振动检测

周 勇^{1,2}, 杨学志^{2,3}, 臧宗迪^{1,2}, 王金诚¹, 吴克伟^{1,2}

(1. 合肥工业大学 计算机与信息学院, 安徽 合肥 230601; 2. 工业安全与应急技术安徽省重点实验室, 安徽 合肥 230601; 3. 合肥工业大学 软件学院, 安徽 合肥 230601)

摘 要:针对无人机平台视觉振动检测不准确的问题,文章提出一种抗无人机晃动的视觉振动检测算法,用于实现无人机拍摄镜头不稳定情况下的振动检测。该算法利用目标振动的局部性和无人机晃动的全局性,从源视频的混合信号中提取出目标振动。首先根据像素点亮度变化的剧烈程度实现目标振动区域与背景区域的分离;其次对图像的背景区域进行特征点提取,并求解背景区域的运动矩阵;然后对目标区域进行运动矫正,去除其包含的背景信号分量,即晃动信号分量;最后提取目标振动区域的局部相位序列,构建振动信号,并检测其振动频率。为了验证该算法的准确性,使用无人机在不同晃动幅度下拍摄多组视频并检测其振动频率。实验结果表明,该算法具有较高的准确性与鲁棒性,能够有效去除无人机晃动干扰,准确提取出视频中的目标振动信号。

关键词:视觉测量;振动检测;无人机;可控金字塔

中图分类号:TP391 **文献标志码:**A **文章编号:**1003-5060(2023)02-0180-09

Phase-based visual vibration detection against UAV shaking

ZHOU Yong^{1,2}, YANG Xuezhi^{2,3}, ZANG Zongdi^{1,2}, WANG Jincheng¹, WU Kewei^{1,2}

(1. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China; 2. Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei 230601, China; 3. School of Software, Hefei University of Technology, Hefei 230601, China)

Abstract: Aiming at the problem of inaccurate visual vibration detection of the unmanned aerial vehicle (UAV) platform, a vision-based vibration detection algorithm against the shaking of UAV is proposed to realize the vibration detection under the condition of unstable lens. The algorithm utilizes the locality of target vibration and the globality of the motion of UAV to extract the target vibration from the mixed signals of source video. Firstly, the vibration region of the target is separated from the background region according to the variation of the pixel brightness. Then the feature points are extracted from the background region and the motion matrix of it is solved. After that, the target region is corrected to remove the signal component of background, which is essentially the motion of UAV. Finally, the local phase sequence of the target vibration region is extracted and the vibration signal is constructed, and then the vibration frequency is detected. In order to verify the accuracy of the algorithm, a UAV is used to shoot multiple sets of vibration videos under different shaking amplitudes and detect their frequencies. Experimental results show that the proposed algorithm has high accuracy and robustness, and can effectively remove the shaking interference of UAV and accurately extract target vibration signals from video.

Key words: vision-based measurement; vibration detection; unmanned aerial vehicle (UAV); steerable pyramid

收稿日期:2021-11-18;修回日期:2022-01-25

基金项目:安徽省科技重大专项资助项目(201903c08020010);安徽高校协同创新资助项目(GXXT-2019-003)

作者简介:周 勇(1997—),男,贵州遵义人,合肥工业大学硕士生;

杨学志(1970—),男,安徽合肥人,博士,合肥工业大学教授,博士生导师,通信作者,E-mail:xyang@hfut.edu.cn.

0 引言

物体的振动特征能够反映该物体的某些本质属性。比如,桥梁等大型建筑的振动特征反映其结构健康状况^[1-2],螺栓等连接件的振动参数反映出工厂设备的运行状态^[3],人体呼吸的振动特征反映了人类的情绪和健康状况^[4]。基于视觉的振动检测方法具有便捷与低成本的优势,且能够避免接触式传感器对目标振动产生的干扰,受到国内外学者的广泛关注^[5]。

视觉振动检测系统利用相机捕捉目标振动视频,将三维图像信号转换为振动信号,进而获取振动频率等参数。固定相机能够完成小范围、地面位置的目标振动测量,然而对于桥梁等大型基础设施的振动检测,固定位置的视觉测量系统则难以实现。这种情况下,通常需要使用多个相机布置测点,相机的部署位置、测量视角等都需要精确计算;此外,实际环境中某些位置并不适合架设相机。

无人机技术的发展为这一问题提供了可能的解决方案,通过无人机搭载相机进行拍摄,能够克服相机部署位置的局限性。然而,无人机在飞行过程中受到气流的影响会产生随机晃动,造成所拍画面不稳定,无人机平台的视觉振动检测系统需要克服这一问题。

对于结构的振动变化,近年来有研究将欧拉视角下的相位分析法应用于振动检测。文献[6-7]将基于相位的运动分析法与欧拉视角相结合,实现了全场运动估计。图像的相位信息相对于亮度信息,对环境光照变化和图像噪声表现出更高的鲁棒性。文献[8-9]使用基于相位的振动提取方法,通过高速摄像机测量悬臂梁、金属杆以及民用基础设施的微小振动,并开展大量的实验验证了该方法的可行性;文献[10]采用基于相位的振动信号提取算法测量了金属杆等材料的模态频率;文献[11]使用基于相位的运动估计和放大方法测量了风力机叶片的谐振频率;文献[12]在基于相位的运动处理框架上实现目标的振动模态分解,并在实验室搭建的小型建筑结构上进行了验证;文献[13]对原有的相位振动检测方法进行了改进,简化了振动波形的提取过程。然而,上述研究所提出的算法只适用于稳定的相机,当应用于无人机平台时,会因气流所引发的相机晃动而失效。

对于无人机平台的振动检测,有学者使用特

征点追踪的方式实现。文献[14-15]利用无人机搭载相机拍摄实验室桥梁模型,检测图像特征点并求解运动矩阵,进而测量目标的振动时程、频率等参数。这两项研究均假设图像背景区域的运动信号反映无人机晃动,并通过手动选取背景特征点提取干扰信号,完成运动校正,从而准确测量目标振动。

但是,特征点追踪的运动分析方法本质上基于亮度信号,对光照变化和图像噪声较为敏感;并且这两项研究中,背景特征点依赖于手工选取,引入了很大的工作量。文献[16]使用无人机捕捉地面人体的呼吸振动信号,首先对无人机所拍摄的视频进行稳像处理后,然后通过基于相位的运动信号提取方式完成呼吸振动检测。然而,稳像算法可能会抑制目标振动信号本身,从而影响检测结果的准确性。

本文提出一种抗无人机晃动的视觉振动检测算法。该算法基于局部相位的信号提取方式,利用晃动信号的全局性和目标振动信号的局部性对信号进行分离。首先根据像素点亮度变化剧烈程度实现目标振动区域与背景区域的分离;其次对图像的背景区域进行特征点提取,并求解背景区域的运动矩阵;然后对目标区域进行运动矫正,去除其包含的背景信号分量,即晃动信号分量;然后提取目标振动区域的局部相位序列,构建振动信号,并检测其振动频率。实验结果表明,该算法能够有效去除无人机晃动干扰,实现无人机在晃动情况下的振动检测。

1 本文算法

本文提出的无人机晃动下的视觉振动检测算法以局部相位提取振动信号为基础,以相位的运动信号提取方式处理振动信号,根据振动区域亮度变化更为剧烈这一特征,定位出振动区域;然后对图像背景区域进行特征点提取,得出背景点的运动矩阵,即无人机在气流等影响下产生的晃动信号的运动矩阵,通过这个运动矩阵对目标区域进行矫正,消除无人机晃动的影响;最后提取目标区域的局部相位序列,构建振动信号并检测目标区域的振动频率。

本文所提算法的流程框图如图1所示,包括3个主要步骤:

- (1) 振动目标定位与分离;
- (2) 无人机晃动信号矫正;
- (3) 基于相位的振动信号提取。

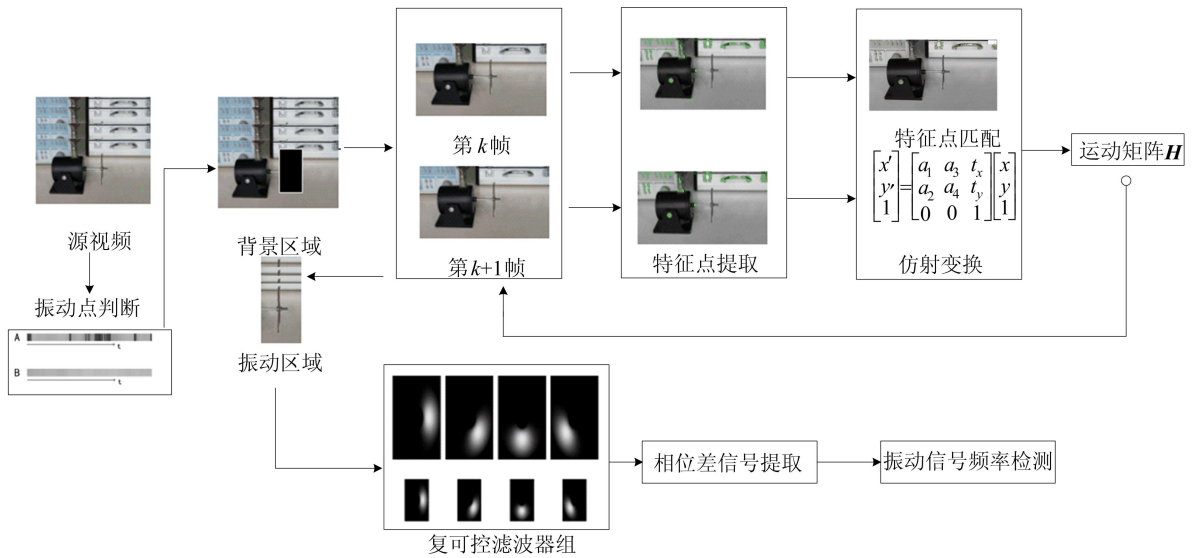


图 1 本文算法流程

1.1 振动目标定位与分离

分离目标区域与背景区域,一方面是为了增强所提取振动信号的纯净度,另一方面是为了利用背景特征点对无人机晃动信号建模。像素点的位置变化可以与其亮度变化相对应,即图像的运动信息可以由其亮度信息来表示,因此本文根据像素点亮度变化特征识别目标点与背景点。目标像素点的运动由目标自身振动引起,而背景像素点的运动则对应无人机的晃动。本文做以下 2 个假设:① 无人机自身的晃动信号,相比较于振动信号,在时间域上是线性的;② 振动信号包含高阶运动分量,不具备线性变化特征。

1.1.1 二阶差分滤波

使用二阶差分滤波的方式可实现振动点与背景点的分离。 I 表示视频的第 1 帧图像; (x, y) 为图像任意一点坐标; $i(x, y, k)$ 为视频第 k 帧图像 (x, y) 点所对应的亮度信号, $1 \leq k \leq N$, N 为视频总帧数,则滤波后信号 $i_f(x, y, k)$ 可表示为:

$$i_f(x, y, k) = i(x, y, k + 2) - 2i(x, y, k + 1) + i(x, y, k) \quad (1)$$

目标点的亮度-时间序列通过该滤波器后,仍有较大的能量剩余;背景点的亮度-时间序列经过二阶差分滤波后则大幅衰减。

实验时选取的不同位置如图 2 所示,图 2 中: A 点代表振动区域的点; B、C 点代表非振动区域的点。A、B、C 3 点的亮度序列 (intensity sequence, IS) 及其滤波后信号的亮度序列 (IS-Filtered) 如图 3 所示。由图 3a、图 3b 可知, A 点在滤波前由振动信号与无人机晃动信号的叠加,在亮度上变化比较明显,亮度-时间序列变化比较剧

烈,在经过二阶差分滤波后还存在大量的能量残余。由图 3c、图 3d 可知, B 点只有无人机晃动信号,因此 B 点属于平坦区域, B 点的亮度-时间序列趋于一个常量,并且变化并不明显,经过二阶差分滤波之后可以看出几乎没有能量残余。由图 3e、图 3f 可知, C 点处于导线边缘,在只有无人机晃动的信号下才会有亮度变化的较大波动。

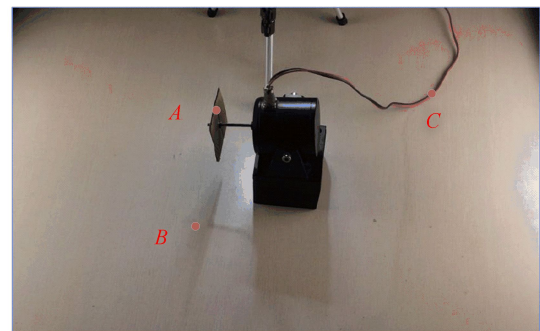
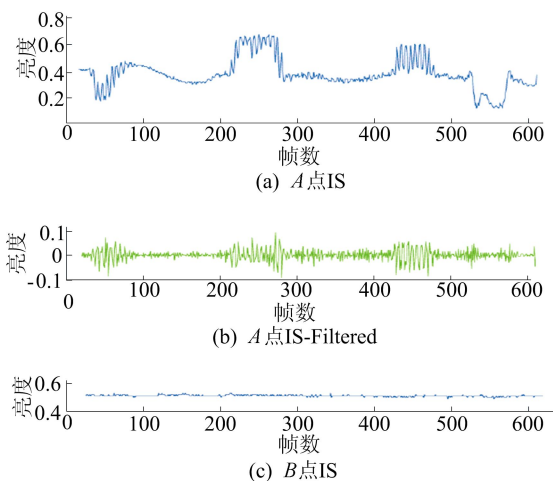


图 2 振动区域与非振动区域



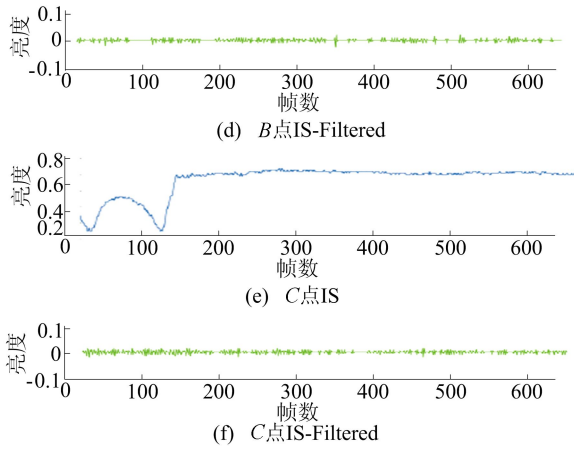


图 3 3 点滤波前后亮度序列

从图 3e、图 3f 还可以看出,通过二阶差分滤波器的亮度-序列趋于 0,由此可知,二阶差分滤波对于滤除线性运动,找寻有振动的区域是有效的。

1.1.2 基于能量的目标分离

在对图像进行二阶差分滤波后,用 $e_f(x, y)$ 表示某一位置 (x, y) 处滤波后亮度信号的能量大小,则有:

$$e_f(x, y) = \sum_{k=1}^N i_f^2(x, y, k) \quad (2)$$

根据整幅图像滤波之后的能量大小,构造振动能量矩阵 \mathbf{E} :

$$\mathbf{E} = \begin{bmatrix} e_f(x_0, y_0) & e_f(x_1, y_0) & \cdots & e_f(x_N, y_0) \\ e_f(x_0, y_1) & e_f(x_1, y_1) & \cdots & e_f(x_N, y_1) \\ \vdots & \vdots & & \vdots \\ e_f(x_0, y_M) & e_f(x_1, y_M) & \cdots & e_f(x_N, y_M) \end{bmatrix} \quad (3)$$

矩阵 \mathbf{E} 表示图像中每个位置所含振动信号的强弱,然后以 σ 为阈值实现 $e_f(x, y)$ 的二值化:

$$E_{BW}(x, y) = f(e_f(x, y)) = \begin{cases} 0, & e_f(x, y) < \sigma; \\ 1, & e_f(x, y) \geq \sigma \end{cases} \quad (4)$$

通过(4)式实现矩阵 \mathbf{E} 的二值化 $\mathbf{E}_{BW} = f(\mathbf{E})$ 。 σ 的选取依赖于经验,一般取 $\sigma = 0.2 \max(e_f(x, y))$ 。

完成以上操作后,二值图像中可能包含振动信号的点会被置 1,背景点则会被置 0。采用图像形态学操作,去除 \mathbf{E}_{BW} 中面积过小的连通区域,得到修饰后的振动能量矩阵 \mathbf{E}'_{BW} 。一方面由于在视频画面存在晃动时,振动目标通常会扫过一块较大的面积,另一方面为了便于后续处理,将 \mathbf{E}'_{BW} 的非零区域扩展为小矩形块,从而得到表征振动目标的模板矩阵 \mathbf{M} 。将模板矩阵 \mathbf{M} 与视频

的每帧图像 \mathbf{I} 相乘即可分离出目标振动区域与背景区域。具体如图 4 所示。

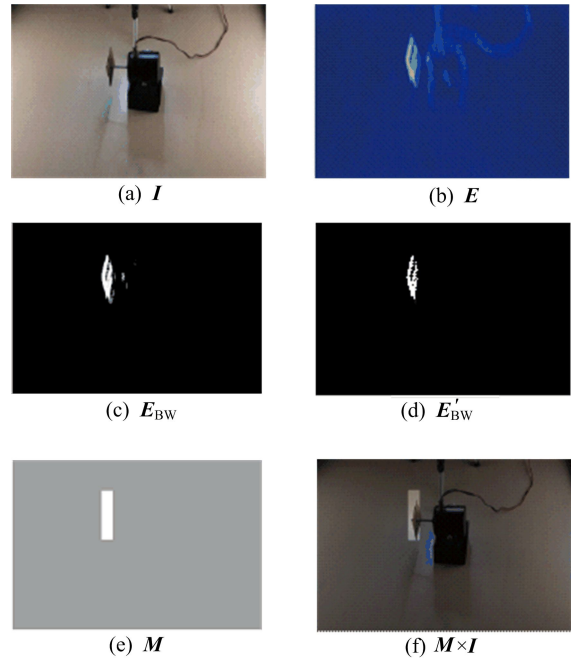


图 4 振动目标定位

1.2 目标区域晃动信号校正

无人机晃动引发图像背景产生整体运动。对图像背景区域进行运动特征的分析则能够实现对无人机晃动信号的估计,处理过程使用角点检测算法(features from accelerated segment test, FAST)^[17]与特征点提取算法(fast retina key-point, FREAK)^[18]算法实现背景区域特征点的检测和匹配,通过采样一致估计(m-estimate sample consensus, MSAC)^[19]剔除误匹配点。完成全局运动模型的建立并构建得到仿射变换矩阵,然后对目标区域进行校正。具体实现方法如下。

1.2.1 特征点提取

根据目标点与周围点的灰度差来确定特征点,FAST 算法步骤如下:① 从图片中选取一个目标像素点 p ,设其像素值为 I_p ;② 设定一个合适的阈值 t ;③ 选取 16 个以目标像素点 p 为中心,半径为 3 的像素点,像素值分别记为 $I_j, j=1, 2, \dots, 16$;④ 满足判决函数有连续的 9 个像素点的那个像素点即为角点。判决函数为:

$$S = |I - I_j| > t \quad (5)$$

1.2.2 特征描述与匹配

FREAK 算法^[18]的采样结构为 6、6、6、6、6、6、6、1,其中 6 代表每层中有 6 个采样点并且这 6 个采样点在一个同心圆上,一共有 7 个同心圆,

最后的 1 表示特征点。采用二进制编码描述特征点,用 F 为编码特征; P_a 为采样点对; N_b 为期望的二进制编码长度, $I(P_a^{r_1})$ 表示采样点对 P_a 中前一个采样点的像素值, $I(P_a^{r_2})$ 表示后一个采样点的像素值,则:

$$F = \sum_{0 \leq a \leq N_b} 2^a T(P_a); \quad (6)$$

$$T(P_a) = \begin{cases} 1, & I(P_a^{r_1}) > I(P_a^{r_2}); \\ 0, & \text{其他} \end{cases} \quad (7)$$

在特征描述中,得到了 512 bit 的二进制描述符,该描述符的列是方差由高到低的排列,高方差表征模糊信息,低方差表征细节信息。因此,选取前 128 bit(即 16 字节)进行匹配,若 2 个待匹配的特征点前 16 字节距离小于设定的阈值,则再用剩余的比特位进行匹配。最终得到视频任意两帧图像之间的匹配点集。

1.2.3 矫正目标区域

完成背景区域的特征点检测和匹配后,使用 MSAC 算法^[19] 完成全局运动模型的建立。MSAC 可以自动判断出正确的匹配点对(内点),剔除误匹配(外点),并且计算出第 1 幅图像的所有特征点映射到第 2 幅图像的运动模型。该模型由如下一个 3×3 的矩阵描述,即

$$\mathbf{H}_i = \begin{bmatrix} a_1 & a_3 & 0 \\ a_2 & a_4 & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad (8)$$

其中: (t_x, t_y) 为平移量; $a_m (m=1, 2, 3, 4)$ 为控制图像的旋转和缩放。为了简化计算,使用以下只有 1 个参数的变换矩阵来代替仿射变换矩阵:

$$\mathbf{H}_i = \begin{bmatrix} s \cos \theta & -s \sin \theta & 0 \\ s \sin \theta & s \cos \theta & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad (9)$$

对 \mathbf{H}_i 求逆,可以消除两帧图像之间背景区域的整体运动。用 \mathbf{G}_k 表示第 1 帧图像背景区域和第 k 帧图像背景区域之间的映射关系,则 \mathbf{G}_k 可以由上述帧与帧之间的变换矩阵 \mathbf{H}_i 累积表示如下:

$$\mathbf{G}_k = \prod_{i=0}^{k-1} \mathbf{H}_i \quad (10)$$

所求 \mathbf{G}_k 描述了背景区域的运动,即实现了无人机晃动信号的估计。 \mathbf{G}_k 可用来对目标振动区域进行校正,去除其包含的无人机晃动分量,保留振动分量。设 \mathbf{O}_k 表示第 k 帧目标振动区域, (x_k, y_k) 为其任意一点, $(x_k, y_k) \in \mathbf{O}_k$, \mathbf{T} 表示校正变换,则有:

$$\begin{bmatrix} x_k' \\ y_k' \\ 1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} x_k \\ y_k \\ 1 \end{bmatrix} = \mathbf{G}_k^{-1} \begin{bmatrix} x_k \\ y_k \\ 1 \end{bmatrix} \quad (11)$$

其中, (x_k', y_k') 为校正后的点坐标。将目标区域的所有点完成校正,得到 $\mathbf{O}_k' = \mathbf{T}(\mathbf{O}_k)$ 。

1.3 基于相位的振动信号提取

完成对目标区域晃动信号矫正后,需要提取振动信号,提取流程如图 1 所示。首先在空间上,使用复数可控滤波器组对视频的每一帧进行不同尺度、不同方向的子带分解,获取到相位变化序列;然后在时域上滤除相位序列的直流分量,滤波后的相位序列即为目标振动信号;对振动信号进行傅里叶变换,检测其幅频响应函数峰值,即可得到振动频率,算法原理如下。

傅里叶变换获取到的空间全局相位与图像的整体运动相关,与之类似,空间局部相位可以用来计算局部运动。复可控金字塔(complex steerable pyramid, CSP)可以实现空间局部相位的提取。本文以一维全局傅里叶基为例说明振动与相位之间的关系,用 $f(x)$ 表示位置 x 处的图像亮度,进行傅里叶级数展开, ω 表示每个子带的空间频率,则有:

$$f(x) = \sum_{\omega=-\infty}^{\infty} A_{\omega} e^{i\omega x} \quad (12)$$

用 $v(t)$ 表示振动引发的位置变化,则有:

$$f(x+v(t)) = \sum_{\omega=-\infty}^{\infty} A_{\omega} e^{i\omega(x+v(t))} \quad (13)$$

其中每个子带可以表示为:

$$S_{\omega}(x, t) = A_{\omega} e^{i\omega(x+v(t))} \quad (14)$$

由于每个子带都是复指数函数,相位 $\omega(x+v(t))$ 含有运动信息。滤除其直流分量,得到相位变化信号 $V_{\omega}(x, t) = \omega v(t)$, $V_{\omega}(x, t)$ 为位置 x 处的振动信号。

复可控金字塔可以将视频的每一帧亮度信号分解为不同尺度、不同方向的局部相位,这些处理后的局部相位信号可提取出视频中的目标振动。用 $C_{\lambda, \theta}$ 表示某一空间尺度 λ 、某一方向 θ 的复可控滤波器,则有:

$$A(x, y, k) e^{i\phi(x, y, k)} = C_{\lambda, \theta}(x, y) \otimes T(\mathbf{O}_k(x, y)) \quad (15)$$

其中: $A(x, y, k)$ 为复可控滤波得到的局部幅值; $\phi(x, y, k)$ 为复可控滤波得到的局部相位。滤除 $\phi(x, y, k)$ 的直流分量,即可得到振动信号 $V(x, y, k)$ 。经傅里叶处理后,可得峰值振动频率 f 为:

$$f = \max_{A(f)} \text{FFT}_k(V(x, y, k)) \quad (16)$$

其中: $A(f)$ 为变换后得到的幅度求最大值; FFT_k 表示对变量 k (即时间) 做傅里叶变换。

2 实验结果与分析

2.1 实验设备及数据

实验系统由 1 台 DJI Mavic Air 2 无人机、1 台 SA-SG030 扫频信号发生器和 1 台 SA-JZ002 模态激振器组成。使用信号发生器模拟真实场景中电机等设备的不同振动频率, 使用无人机在不同晃动幅度下拍摄激振器视频, 测量激振器顶端杆的振动频率。无人机所拍视频的分辨率为 1 080 像素 \times 1 920 像素, 帧率为 60 帧/s。通过手动操纵遥控杆可使无人机产生 3 种不同幅度(小幅、中幅、大幅)的晃动, 其中小幅晃动由无人机静止拍摄产生。

使用无人机拍摄激振器振动视频, 拍摄时长为 10 s。计算多个特征点在不同帧的坐标间的欧式距离, 选择最大距离作为无人机晃动幅度的估计。按照无人机晃动程度的不同将实验视频分为 3 组: 组别 1(晃动幅度 0~30 像素)、组别 2(晃动幅度 30~60 像素)、组别 3(晃动幅度 60~90 像素)。激振器的振动频率可通过信号发生器调节, 将激振器频率分别设置为 5.2、10.2、15.2 Hz, 每个频率点拍摄 10 段视频, 一共采 90 段视频作为实验数据集。实验场景如图 5 所示。



图 5 实验场景

2.2 评价标准

(1) 平均绝对误差 M_e , 用来反映所测频率与真实频率的绝对误差平均值, 计算公式为:

$$M_e = \frac{1}{N_v} \sum_{n=1}^{N_v} |f^n - f_{tr}^n| \quad (17)$$

其中: N_v 为该组实验视频的数量; f^n 为序号为 n 的视频测得的频率值; f_{tr}^n 为序号为 n 的视频的真实频率值。

(2) 均方误差 R_{MSE} , 用来反映在多次测量结果中, 检测频率偏离真实频率的程度, 计算公式为:

$$R_{MSE} = \sqrt{\frac{1}{N_v} \sum_{n=1}^{N_v} (f^n - f_{tr}^n)^2} \quad (18)$$

2.3 实验结果

2.3.1 验证实验

本文使用对真实频率为 10.2 Hz 的激振器进行频率检测, 从不同晃动幅度的 3 组视频中各选取 1 段视频进行处理, 生成的频谱图如图 6 所示。

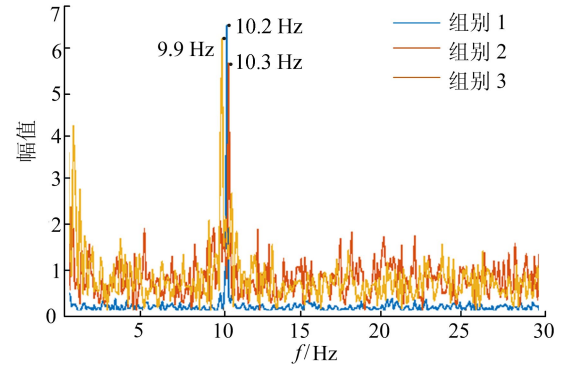


图 6 不同晃动幅度检测频谱对比

从图 6 可以看出, 组别 1 测得的峰值频率为 10.2 Hz, 组别 2 为 10.3 Hz, 组别 3 为 9.9 Hz。结果表明, 本文算法在不同晃动幅度下均能准确测得振动信号, 当晃动幅度增大时, 误差相应增大。为了进一步验证本文算法的检测效果, 计算 3 种晃动幅度、3 种振动频率下的平均绝对误差和均方误差, 对采集的全部 90 段视频进行频率检测, 结果见表 1 所列。

表 1 本文算法检测结果

组别	f/Hz	M_e/Hz	R_{MSE}/Hz
组别 1	5.2	0.090	0.105
	10.2	0.100	0.127
	15.2	0.070	0.095
组别 2	5.2	0.130	0.182
	10.2	0.110	0.145
	15.2	0.120	0.141
组别 3	5.2	0.180	0.219
	10.2	0.150	0.207
	15.2	0.140	0.167

由表 1 可知本文算法的平均绝对误差控制在 0.2 Hz 以内, 均方误差不超过 0.22 Hz。

2.3.2 对比实验

将本文算法与文献[13]使用的相位检测算法、文献[15]使用的特征点追踪算法、文献[16]使用的稳像后相位提取算法进行频率检测对比。组别 1 中振动频率为 10.2 Hz 的视频在各方法下的频率检测结果如图 7 所示。从图 7 可以看出, 文

献[13]、文献[15]、文献[16]算法的峰值频率分别对应 0.3、10.6、8.9 Hz,而本文算法检测结果为 10.2 Hz。相比于本文算法,其他 3 种算法都存在较大误差。由该组实验可以看出,本文算法能够有效去除无人机晃动干扰,准确提取出视频中的振动信号,相比于其他算法的检测结果,本文算法的误差更小。为了进一步对比上述算法的检测效果,对采集的全部 90 段视频进行频率检测,检测结果见表 2~表 4 所列。

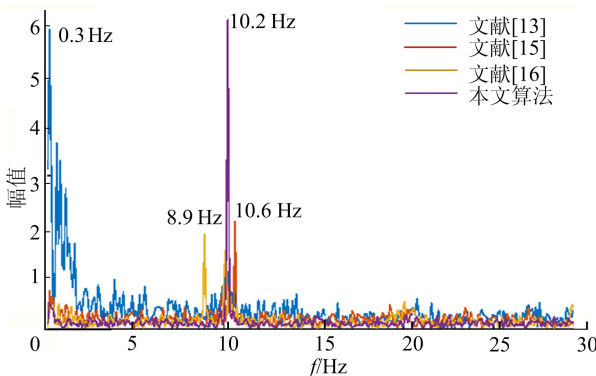


图 7 10.2 Hz 下不同算法频率检测对比

表 2 0~30 像素晃动下检测结果

f/Hz	算法	M_e/Hz	R_{MSE}/Hz
5.2	文献[13]	3.010	3.310
	文献[15]	0.320	0.379
	文献[16]	0.340	0.417
	本文算法	0.100	0.118
10.2	文献[13]	5.830	8.503
	文献[15]	0.240	0.303
	文献[16]	1.140	0.416
	本文算法	0.080	0.100
15.2	文献[13]	7.890	9.293
	文献[15]	0.220	0.316
	文献[16]	1.210	1.383
	本文算法	0.060	0.089

表 3 30~60 像素晃动下检测结果

f/Hz	算法	M_e/Hz	R_{MSE}/Hz
5.2	文献[13]	2.520	2.179
	文献[15]	0.410	0.699
	文献[16]	0.520	0.699
	本文算法	0.090	0.114
10.2	文献[13]	4.040	5.454
	文献[15]	0.270	0.333
	文献[16]	0.860	1.029
	本文算法	0.120	0.141
15.2	文献[13]	6.810	7.719
	文献[15]	0.250	0.311
	文献[16]	1.440	1.742
	本文算法	0.120	0.148

表 4 60~90 像素晃动下检测结果

f/Hz	算法	M_e/Hz	R_{MSE}/Hz
5.2	文献[13]	3.060	3.567
	文献[15]	0.650	0.786
	文献[16]	0.670	0.870
10.2	本文算法	0.300	0.382
	文献[13]	4.140	4.678
	文献[15]	0.370	0.468
	文献[16]	1.160	1.920
15.2	本文算法	0.130	0.170
	文献[13]	8.170	11.332
	文献[15]	0.360	0.434
	文献[16]	3.050	3.716
	本文算法	0.190	0.226

从表 2~表 4 可以看出,对比文献[13]、文献[15]、文献[16]算法,分别在 5.2、10.2、15.2 Hz 下对应 3 个不同幅度无人机晃动的检测结果,发现本文算法均优于另外 3 种算法。

2.3.3 实验分析

本实验对本文算法与文献[13]、文献[15]、文献[16]3 种算法进行准确性分析,结果见表 5 所列。

表 5 准确性分析

算法	总体 M_e/Hz	总体 R_{MSE}/Hz
文献[13]	5.052	6.886
文献[15]	0.343	0.446
文献[16]	1.154	1.730
本文算法	0.132	0.186

从表 5 可以看出,本文算法取得了最佳的检测效果,分组平均绝对误差和均方误差均保持在 0.2 Hz 的范围内,总体平均绝对误差为 0.132;文献[15]算法检测性能次之,总体平均绝对误差 0.343;文献[16]算法整体检测性能欠佳,总体平均绝对误差超过了 1;而文献[13]算法整体误差过大,超过了可接受的范围,因此可判定该算法对于无人机晃动下的振动检测失效。

原因在于:文献[13]不加区分地将全场运动信号转化为相位信息,未处理相机晃动,引入了大量低频干扰,导致目标振动信号被湮没;文献[15]算法通过手动选定背景特征点实现运动矫正,保留了振动信号,但由于其采用基于亮度特征的振动提取方式,因而其振动检测准确性上低于本文的相位检测法;文献[16]算法虽然针对无人机引发的相机晃动进行了稳像处理,但是由于其未区分目标振动信号和晃动信号,从而导致振动信号

产生了畸变。

本文算法实现了振动和背景区域的分离,以背景区域的特征点完成对无人机运动矩阵的求解,从而去除振动区域的晃动干扰,准确提取出了目标振动视频中的待测振动信号。

在不同晃动幅度下将本文算法与其他 3 种算法的平均绝对误差进行对比,见表 6 所列。从表 6 可以看出:对于文献[15]、文献[16]以及本文算法,随着无人机晃动幅度的增大,3 种算法的频率检测误差均明显增大;文献[15]采用特征点追踪框架,当晃动增大时,其特征点匹配错误的可能性增加,造成性能下降;文献[16]使用稳像算法对视频进行处理后,再进行基于相位的振动检测,由于视频稳像依赖特征点匹配,因而晃动幅度增加会影响其稳像预处理的性能,进而降低频率检测准确率。

对于本文算法,由于无人机运动矩阵的求解同样依赖特征点匹配,检测准确率必然随着晃动幅度的增大而降低。

表 6 3 种晃动幅度下平均绝对误差分析 单位:Hz

算法	组别 1	组别 2	组别 3
文献[13]	5.577	9.810	5.123
文献[15]	0.260	0.310	0.460
文献[16]	0.533	0.940	1.627
本文算法	0.080	0.110	0.207

在不同频率下将本文算法与其他 3 种算法的平均绝对误差进行对比,结果见表 7 所列。

表 7 不同振动频率下平均绝对误差分析 单位:Hz

f	5.2	10.2	15.2
文献[13]	2.863	4.670	7.623
文献[15]	0.460	0.293	0.277
文献[16]	0.510	1.053	1.900
本文算法	0.163	0.110	0.123

由表 7 可知,由于文献[13]算法在实验场景下已基本失效,因此不做进一步分析;文献[15]算法和本文算法在不同振动频率下的检测性能无明显差别;而文献[16]算法则在测量高频信号时,准确率明显降低。原因在于:文献[16]算法基于视频的全局稳像,稳像过程会去除剧烈变化信号,保留图像序列的平滑运动,如果目标的振动频率较高,那么算法可能会对目标本身的振动产生抑制效果,进而影响频率检测的准确性;文献[15]算法通过人工选择背景特征点实现晃动信号的估计;

本文算法则在自动分离背景区域和振动区域的基础上,使用背景区域的特征点求解无人机运动矩阵。因此这 2 种算法均不会对目标本身的振动产生影响。

2.3.4 算法时间复杂度分析

将本文算法时间复杂度与文献[13]、文献[15]、文献[16]算法时间复杂度进行对比分析,结果见表 8 所列,视频每帧分辨率为 1 080 像素 \times 1 920 像素,每个视频总共为 600 帧,使用 5 个不同的视频进行实验。

从表 8 可以看出,本文算法的时间复杂度优于文献[13]、文献[16],但比文献[15]更耗时。这是由于文献[13]使用的是全局相位检测算法,由于对视频每帧整体进行计算,本文算法在矫正分离后只对振动区域进行相位检测算法,只计算每一帧部分位置,因而本文算法耗时更少。

对比文献[15]采用特征点追踪算法,在算法的时间复杂度上比本文算法更优。文献[16]采用全局稳像再使用相位检测算法,首先稳像是针对全局进行,而本文算法的矫正只针对目标区域,在这一阶段本文的算法就会耗时更少。而且文献[16]在使用稳像算法之后再再进行全局的相位检测算法,整体来看,本文算法的时间复杂度明显优于文献[16]。

表 8 不同算法时间性能分析 单位:s

视频序号	文献[13]	文献[15]	文献[16]	本文算法
1	151.35	91.17	229.02	107.91
2	150.56	90.98	230.01	106.36
3	151.43	91.03	230.13	106.23
4	151.69	90.12	231.15	108.55
5	151.02	90.96	230.78	107.21

3 结 论

本文提出了一种抗无人机晃动的视觉振动检测算法,该算法利用目标物体的振动特征,从视频中分离出振动目标和背景区域,利用背景区域计算出无人机在气流扰动下晃动的运动矩阵,利用得到的运动矩阵实现对目标区域里叠加的无人机晃动信号进行矫正,得到较为纯净的振动信号,然后通过复可控滤波器提取振动信号的相位序列,检测其振动频率。

在真实使用场景中,振动目标不一定总能够保持在拍摄画面内,当无人机晃动幅度过大时,目标可能会短暂的丢失在视频帧中,这会影响到本

文算法的准确性。

在未来的工作中,将针对更大幅度的无人机晃动以及更为复杂的无人机测量场景开展研究,以拓宽本文算法的应用范围。

[参 考 文 献]

- [1] BROWNJOHN J, STEFANO A D, XU Y L, et al. Vibration-based monitoring of civil infrastructure: challenges and successes[J]. *Journal of Civil Structural Health Monitoring*, 2011, 1(3/4): 79-95.
- [2] 蒋新花, 丁德宝, 沈庆宏. 应用于桥梁结构健康监测的低噪声振动测量系统实现[J]. *电子测量技术*, 2014(2): 116-119.
- [3] SHENGCHANG J, YONGFEN L, YANMING L. Research on extraction technique of transformer core fundamental frequency vibration based on OLCM[J]. *IEEE Transactions on Power Delivery*, 2006, 21(4): 1981-1988.
- [4] CHAN A M, FERDOSI N, NARASIMHAN R. Ambulatory respiratory rate detection using ECG and a triaxial accelerometer[C]//Conference of the IEEE Engineering in Medicine and Biology Society. [S. l.]: IEEE, 2013: 4058-4061.
- [5] AOYAMA T, LI L, JIANG M, et al. Vibration sensing of a bridge model using a multithread active vision system[J]. *IEEE/ASME Transactions on Mechatronics*, 2018, 23(1): 179-189.
- [6] FLEET D J, JEPSON A D. Computation of component image velocity from local phase information[J]. *International Journal of Computer Vision*, 1990, 5(1): 77-104.
- [7] WADHWA N, RUBINSTEIN M, DURAND F, et al. Phase-based video motion processing[J]. *ACM Transactions on Graphics (TOG)*, 2013, 32(4): 1-10.
- [8] CHEN J G, WADHWA N, CHA Y J, et al. Modal identification of simple structures with high-speed video using motion magnification[J]. *Journal of Sound and Vibration*, 2015, 345(9): 58-71.
- [9] WADHWA N, CHEN J G, SELLOON J B, et al. Motion microscopy for visualizing and quantifying small motions[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(44): 11639-11644.
- [10] DAVIS A, DOUMAN K L, CHEN J G, et al. Visual vibrometry: estimating material properties from small motions in video[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 732-745.
- [11] SARRAFI A, MAO Z, NIEZRECKI C, et al. Vibration-based damage detection in wind turbine blades using phase-based motion estimation and motion magnification[J]. *Journal of Sound & Vibration*, 2018, 421(12): 300-318.
- [12] YANG Y, DORN C, MANCINI T, et al. Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification[J]. *Mechanical Systems and Signal Processing*, 2017, 85(15): 567-590.
- [13] PENG C, ZENG C, WANG Y. Camera-based micro-vibration measurement for lightweight structure using an improved phase-based motion extraction[J]. *IEEE Sensors Journal*, 2019, 20(5): 2590-2599.
- [14] YOON H, SHIN J, SPENCER B F. Structural displacement measurement using an unmanned aerial system[J]. *Computer-aided Civil & Infrastructure Engineering*, 2018, 33(3): 183-192.
- [15] TK A, TAN B, HD B, et al. Swaying displacement measurement for structural monitoring using computer vision and an unmanned aerial vehicle[J]. *Measurement*, 2020, 159(15): 107769.
- [16] PERERA A G, KHANAM F, AL-NAJI A A, et al. Detection and localisation of life signs from the air using image registration and spatio-temporal filtering[J]. *Remote Sensing*, 2020, 12(3): 577-1-577-22.
- [17] ROSTEN E, PORTER R, DRUMMOND T. Faster and better: a machine learning approach to corner detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2010, 32(1): 105-119.
- [18] ALAHI A, ORTIZ R, VANDERGHEYNST P. FREAK: fast retina keypoint[C]//IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2012: 510-517.
- [19] TORR P, ZISSERMAN A. MLESAC: a new robust estimator with application to estimating image geometry[J]. *Computer Vision and Image Understanding*, 2000, 78(1): 138-156.

(责任编辑 李 凯)