

基于 MeanShift 算法的光子多普勒测速信号提取方法研究

王 玮 顾国华 何伟基 陈 钱

(南京理工大学江苏省光谱成像与智能感知重点实验室, 江苏 南京 210094)

摘要 针对目前光子多普勒测速(PDV)信号处理方法抗噪性能不足的现状,提出了一种对激光多普勒信号的短时傅里叶变换(STFT)频谱图进行轨迹搜索找出有效目标信号的方法,该方法利用对搜索方向进行加权的 MeanShift 迭代过程来得到密度函数极大值点,然后将该点作为有效信号的轨迹点。实验表明,在给出合适的初始迭代点的情况下,该算法能准确地从含有大量噪声及干扰的短时傅里叶变换频谱图中提取出有效信号,对于强噪声条件下这种 PDV 信号提取算法具有较强的稳健性。

关键词 信号处理; PDV 信号提取; 短时傅里叶变换; MeanShift

中图分类号 TN249 **文献标识码** A **doi**: 10.3788/CJL201441.1108006

Extraction Method of Photon Doppler Velocimetry Signal Based on MeanShift Algorithm

Wang Wei Gu Guohua He Weiji Chen Qian

(*Jiangsu Key Lab of Spectral Imaging & Intelligence Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China*)

Abstract For the situation where the current signal extraction algorithms for the photon Doppler velocimetry (PDV) data are weak in the performance of anti-noise, a method based on trajectory search with the short time Fourier transformation (STFT) spectrum of the PDV data is proposed to extract the valid target signal. The Algorithm uses the iterative MeanShift progress weighting the search direction to get the maximum value points of the density function and these points can be regarded as the trajectory points of valid signal. Experimental results show that, when the proper initial point is given, the algorithm can accurately extract the valid signal from the short-time Fourier transform spectrum which contains a large amount of noise and interference signal, this signal extraction algorithm for PDV has strong robustness in strong noise condition.

Key words signal processing; photon Doppler velocimetry; short time Fourier transformation; MeanShift

OCIS codes 120.3180; 120.7250; 150.1135

1 引言

光子多普勒测速^[1-3](PDV)是一种新型的全光纤激光测速系统,具有非接触测量、时间/速度分辨率高、抗干扰性强等优点,广泛应用于冲击波、爆轰波以及其他短时高速运动的测量应用。检测到的 PDV 信号往往十分微弱,具有样本数量少、信噪比

低、非平稳等特点。目前 PDV 信号处理方法主要有条纹法、短时傅里叶变换(STFT)^[4-5]以及小波变换^[6-7]等,条纹法通过寻找到信号相邻的极大值和极小值对应的时间来得到差频周期,进而得到物体的运动速度,这种方法的精度高,但是靶面角度的改变会引起信号周期的不完整,条纹无法确定不完整

收稿日期: 2014-04-03; **收到修改稿日期**: 2014-06-18

基金项目: 国家自然科学基金青年基金(61101196)、国家自然科学基金面上项目(61271332)

作者简介: 王 玮(1989—),男,硕士研究生,主要从事光电成像和光电探测等方面的研究。

E-mail: wangwei430626@163.com

导师简介: 顾国华(1966—),男,研究员,博士生导师,主要从事光电成像和光电探测等方面的研究。

E-mail: gghnjst@163.com(通信联系人)

周期对应时间内的速度,同时这种方法对高频噪声敏感,抗噪性能不足;STFT 则是通过选取一个固定宽度的滑动窗口,将非平稳信号截取为一段段小样本数量的近平稳的信号,然后分别对截得的信号进行傅里叶变换,得到每段中的频谱,并将此段频谱模极大值所对应的频率作为这段时间的差频频率,进而求得物体的运动速度,这种方法在噪声和干扰信号的强度低于有效信号的强度时有效,而在强噪声干扰的情况下则无法得到有效的速度信号;小波变换虽然采用时-频域变窗口的方法来分析非平稳信号,克服了 STFT 使用固定窗口的缺点,但是在提取差频信号的过程中仍然采用了模极大值的方法,所以仍然无法在强噪声干扰的情况下准确地提取出运动物体的速度曲线。事实上,以上三种 PDV 信号提取方法本质上都是基于信号幅值的判别,默认有效信号的强度高于噪声干扰的强度,这些方法在没有较强的背景信号和噪声的情况下简单有效,但是若原始信号伴随着大量噪声和背景干扰(例如炸弹破片速度),则无法通过上述方法得到有效信号。对此,本文提出了一种基于均值偏移(MeanShift)^[8-9]的邻域迭代对激光多普勒信号^[10-11]的 STFT 频谱图进行轨迹搜索,进而提取有效目标的信号的方法。这种方法在进行信号提取时不仅考虑了信号幅值,而且考虑了信号曲线的几何形态,从而能够准确地将有效信号从含有大量噪声的时频图中分离出来,方便数据的后续处理。

2 传统的基于 STFT 的 PDV 信号处理方法

利用傅里叶变换来获取信号的频谱需要用到信号的全部时域信息,不能反映出信号频率随时间的变化情况,所以傅里叶变换不能处理非平稳信号,鉴于傅里叶变换的缺陷,学者们提出一种利用时间窗将信号分帧处理的思想,这种信号分析方法称为时间-频率分析。而窗固定的时间-频率分析方法即为 STFT。STFT 其主要思想是将信号加窗,将加窗后的信号再进行傅里叶变换,加窗后使得变换为短时间内的局部谱,窗函数可以在信号的整个时间轴平移,因此利用窗函数可以得到任意时间段的局部频谱,从而实现非平稳信号的时间-频率分析,离散信号的 STFT 定义为

$$X_k(\omega) = \sum_{m=-\infty}^{+\infty} x(m)\omega(k-m)\exp(-j\omega m). \quad (1)$$

式中 $x(m)$ 为原始离散信号, $\omega(m)$ 为窗函数,当 k

取不同值时, $\omega(k-m)$ 沿着 $x(m)$ 序列滑动, $X_k(\omega)$ 是信号的 STFT 频谱,可见 $X_k(\omega)$ 同时是时间 k 和频率 ω 的函数。测不准原理指出,不论采用何种函数作为窗函数,其时间窗和频率窗宽度的乘积的最小值都为 2,也就是说,不可能在时间和频率两个空间同时以任意精度逼近被测信号,因此必须根据信号的实际情况选择适当大小的窗函数对信号频谱的时间分辨率和频率分辨率做适当的折衷。

在传统的基于 STFT 的 PDV 信号处理方法中,首先对原始信号进行 STFT 变换得到时间-频率谱,然后在时间-频率谱中的各个时间采样点上提取出能量最大值点作为有效信号的轨迹点。假设原始信号的 STFT 频谱矩阵大小为 $m \times n$, 频谱中的有效信号的轨迹点为 (i, r_i) , 其中, $i \in \{0, 1, \dots, m-1\}$, $r_i \in \{0, 1, \dots, n-1\}$, 那么可以通过如下公式求得 (i, r_i) :

$$\max[X_i(\omega)] = X_i(r_i), \quad i = 0, 1, \dots, m-1. \quad (2)$$

最终频谱矩阵中有效信号的轨迹点集合为 $\{(i, r_i) \mid i = 0, 1, \dots, m-1\}$, 可以看出,基于 STFT 的 PDV 信号处理算法简单易行,但是仅仅基于信号幅值的判别,默认有效信号的强度高于噪声干扰的强度,若被测信号中含有强噪声和背景干扰,那么采用基于 STFT 的 PDV 信号处理算法就不能达到预期效果。

3 基于 MeanShift 算法的 PDV 信号处理方法原理

由于雷管爆炸时测量环境的特殊复杂,PDV 系统采集到的原始信号进行 STFT 变换后得到的时间-频率图含有大量的噪声以及背景干扰,而且干扰噪声的强度比有效信号的强度还要大(如图 2 所示),故传统的 PDV 信号提取算法不再适用。MeanShift 最早用来估计概率密度梯度函数,一般是指一个迭代的步骤,即先算出当前点的偏移均值,移动该点至其偏移均值,然后以此为新的起始点继续移动,直到满足一定的条件结束。本文提出的基于 MeanShift 的邻域迭代对 PDV 信号的 STFT 频谱图进行信号提取的算法的主要原理是:在对原始 PDV 信号进行 STFT 的基础上,对生成的时频矩阵实施强度阈值滤波和二值化处理得到矩阵 $D_{m \times n}$, 然后对 $D_{m \times n}$ 中数据块进行连通域提取,根据轨迹斜率变化不能大于某一阈值的约束条件,给出了由 $D_{m \times n}$ 连通域重心组成的搜索方向约束链表,最后利用

MeanShift 迭代过程总是指向概率密度函数增加最大的方向这一特性来对 $\mathbf{D}_{m \times n}$ 进行有效信号的轨迹搜索,并且引入与轨迹斜率变化相关的权重函数来引导搜索沿着搜索方向约束链表给出的方向进行。

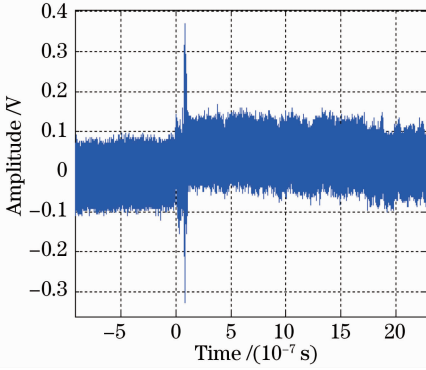


图 1 爆炸冲击波的多普勒信号
Fig.1 PDV signal of shock wave

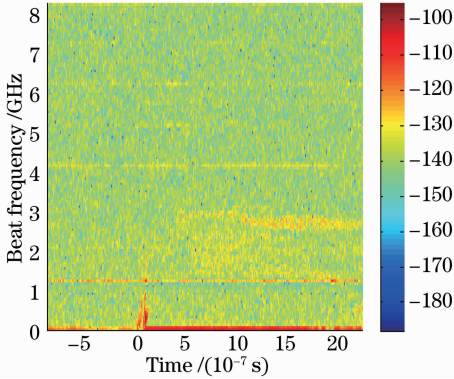


图 2 图 1 所示信号的 STFT 频谱图

Fig.2 STFT spectrum of the signal shown in Fig.1

对一个概率密度函数 $f(p)$,已知 d 维空间中 n 个采样点 $p_i (i = 1, 2, \dots, n)$, $f(p)$ 的核函数估计^[8]为

$$\hat{f}(p) = \frac{\sum_{i=1}^n K\left(\frac{p_i - p}{h}\right) \omega(p_i)}{h^d \sum_{i=1}^n \omega(p_i)}, \quad (3)$$

点集 $\{p_i | i = 1, 2, \dots, n\}$ 对应的 MeanShift 式子^[8]为

$$\mathbf{M}_h(x) = \frac{\sum_{i=1}^n (p - p_i) G\left(\frac{p - p_i}{h}\right) \omega(p_i)}{\sum_{i=1}^n G\left(\frac{p - p_i}{h}\right) \omega(p_i)}, \quad (4)$$

式中 h 为带宽系数, $\omega(p_i) \geq 0$ 是一个赋给采样点 p_i 的权重, $K(r)$ 是一个核函数,并且满足 $\int K(r) dr = 1$ 。

另外定义:核函数 $K(r)$ 的剖面函数^[8] $k(r)$,使得

$$K(r) = k(\|r\|^2), \quad (5)$$

$k(r)$ 的负导函数 $g(r)$,即 $g(r) = -k'(r)$,其对应的核函数

$$G(r) = g(\|r\|^2), \quad (6)$$

概率密度函数 $f(p)$ 的梯度 $\nabla f(p)$ 的估计^[8]为

$$\hat{\nabla} f(p) = \nabla \hat{f}(p) = \frac{2 \sum_{i=1}^n (p - p_i) k'\left(\frac{\|p - p_i\|^2}{h}\right) \omega(p_i)}{h^{d+2} \sum_{i=1}^n \omega(p_i)}. \quad (7)$$

由上面的定义, $g(r) = -k'(r)$, $G(r) = g(\|r\|^2)$, (7) 式可以重写为

$$\hat{\nabla} f(p) = \frac{2 \sum_{i=1}^n (p - p_i) G\left(\frac{p - p_i}{h}\right) \omega(p_i)}{h^{d+2} \sum_{i=1}^n \omega(p_i)} = \frac{2}{h^2} \frac{\sum_{i=1}^n G\left(\frac{p - p_i}{h}\right) \omega(p_i)}{h^d \sum_{i=1}^n \omega(p_i)} \times \frac{\sum_{i=1}^n (p - p_i) G\left(\frac{p - p_i}{h}\right) \omega(p_i)}{\sum_{i=1}^n G\left(\frac{p - p_i}{h}\right) \omega(p_i)}, \quad (8)$$

(8) 式等号右边的第二项就是以 $G(p)$ 为核函数对概率密度函数 $f(p)$ 的估计,记作 $\hat{f}_G(p)$,第三项就是由(4)式所定义的 MeanShift 式子,将(3)式定义的 $\hat{f}(p)$ 重新记作 $\hat{f}_K(p)$,于是

$$\hat{\nabla} f(p) = \nabla \hat{f}_K(p) = \frac{2}{h_2} \hat{f}_G(p) \mathbf{M}_h(p), \quad (9)$$

也即

$$\mathbf{M}_h(p) = \frac{1}{2} h^2 \frac{\nabla \hat{f}_K(p)}{\hat{f}_G(p)}. \quad (10)$$

所以,用核函数 G 在 p 点计算得到的向量 $\mathbf{M}_h(p)$ 正比于归一化的用核函数 K 估计的概率密度的函数 $\hat{f}_K(p)$ 的梯度^[8],归一化因子为用核函数 G 估计的 p 点的概率密度,因此 MeanShift 向量 $\mathbf{M}_h(p)$ 总是指向概率密度增加最大的方向^[8]。

令

$$m_h(p) = \frac{\sum_{i=1}^n G\left(\frac{p_i - p}{h}\right) \omega(p_i) p_i}{\sum_{i=1}^n G\left(\frac{p_i - p}{h}\right) \omega(p_i)}, \quad (11)$$

则

$$\mathbf{M}_h(p) = m_h(p) - p. \quad (12)$$

给定一个初始点 p ,核函数 $G(p)$ 以及容许误差

ϵ , MeanShift 算法循环执行下面三步^[8], 直至满足结束条件:

- 1) 计算 $m_h(p)$;
- 2) 把 $m_h(p)$ 赋给 p ;
- 3) 如果 $\|m_h(p) - p\| < \epsilon$, 结束循环; 若不然, 继续执行步骤 1)。

由于 $m_h(p) = p + \mathbf{M}_h(p)$, 因此上面的步骤也就是不断地沿着概率密度的梯度方向移动。在满足一定条件下, MeanShift 算法一定会收敛到该点附近的峰值, 所以可以用 MeanShift 算法的这种特性来指导有效信号的搜索路径。

为了根据数据的空间分布从宏观上估计搜索轨迹的搜索方向, 宏观上保证搜索沿着正确的方向进行, 需建立一个搜索方向约束链表 CList。

原始信号的 STFT 频谱图 $P_{m \times n}$ 经过阈值滤波和二值化得到时间-频率矩阵 $\mathbf{D}_{m \times n}$, $\mathbf{D}_{m \times n}$ 由大小不一的连通数据块 \mathbf{D}_i , ($i = 1, 2, \dots, k$) 组成, 数据块 \mathbf{D}_i 对应的重心坐标为

$$\begin{cases} x\mathbf{D}_i = \sum_{j=0}^{m_i} x_{i,j}/m_i \\ y\mathbf{D}_i = \sum_{j=0}^{m_i} y_{i,j}/m_i \end{cases}, i = 1, 2, \dots, k, \quad (13)$$

式中点 $(x_{i,j}, y_{i,j})$ 属于连通域 \mathbf{D}_i 的数据点, m_i 表示 \mathbf{D}_i 含有数据点的个数。由点集 $\mathbf{D}_C = \{(x\mathbf{D}_i, y\mathbf{D}_i) | i = 1, 2, \dots, k\}$ 构建的矩阵 \mathbf{MD}_C 是对原先时频矩阵 $\mathbf{D}_{m \times n}$ 的一个简化和近似:

$$\mathbf{MD}_C(i, j) = \begin{cases} 1, & i = \text{round}(xD_t), j = \text{round}(yD_t) \\ 0, & \text{else} \end{cases}, \quad i = 1, 2, \dots, m, j = 1, 2, \dots, n, \quad (14)$$

式中 round 表示四舍五入的函数。

信号轨迹作为一条连续曲线, 相邻两次搜索的方向向量不会变化很大, 假设夹角不会超过角度 ζ , 那么可以基于这点假设对 \mathbf{D}_C 进行搜索生成一条轨迹方向约束链表。假设相邻两次搜索得到的约束链表节点为 P_{i-2}, P_{i-1} , 对于 P_{i-1} 最邻近的点 P_i 若满足

$$\frac{\overrightarrow{P_{i-2}P_{i-1}} \cdot \overrightarrow{P_{i-1}P_i}}{\|\overrightarrow{P_{i-2}P_{i-1}}\| \|\overrightarrow{P_{i-1}P_i}\|} \geq \cos \zeta, \quad (15)$$

$$\text{Distance}(P_{i-1}, P_i) \leq D_{TC}, \quad (16)$$

则将 P_i 加入搜索方向约束链表, 否则删除 P_i 。最终, CList 由有序点列 $P_{i_1}, P_{i_2}, \dots, P_{i_r}$ 构成, i_1, i_2, \dots, i_r 代表连通域的编号, r 表示 CList 链表的长度。

CList 节点按照存放顺序构成的点序列代表了轨迹搜索的大致方向, 规定了各个连通域之间的搜索顺序和大致方向。搜索方向越接近 CList 序列定义的搜索方向, 那么位于这个搜索方向附近的点应该具有更大的权重, 所以可以用搜索方向与 CList 序列定义的搜索方向之间的夹角相关的函数 $\exp(|\cos \alpha|)$ 作为 MeanShift 公式中的权重函数。假设点 $p = (x, y)$ 位于编号为 i_j 的连通域中, CList 中相邻两个节点坐标 $P_{i_{j-1}} = (xD_{i_{j-1}}, yD_{i_{j-1}}), P_{i_j} = (xD_{i_j}, yD_{i_j})$, 则点 p 的权重函数

$$\begin{aligned} \omega(p) &= \exp(|\cos \alpha|) = \exp\left(\left|\frac{\overrightarrow{P_{i_{j-1}}P_{i_j}} \cdot \overrightarrow{P_{i_j}P}}{\|\overrightarrow{P_{i_{j-1}}P_{i_j}}\| \|\overrightarrow{P_{i_j}P}\|}\right|\right) = \\ &= \exp\left(\left|\frac{(x - xD_{i_j})(xD_{i_{j-1}} - xD_{i_j}) + (y - yD_{i_j})(yD_{i_{j-1}} - yD_{i_j})}{\sqrt{(x - xD_{i_j})^2 + (y - yD_{i_j})^2} \sqrt{(xD_{i_{j-1}} - xD_{i_j})^2 + (yD_{i_{j-1}} - yD_{i_j})^2}}\right|\right). \end{aligned} \quad (17)$$

此外, MeanShift 式子中核函数 $G(r)$ 取为高斯核函数,

$$G\left(\frac{p' - p}{h}\right) = \exp\left(-\left\|\frac{p' - p}{h}\right\|^2\right), \quad (18)$$

式中常数 h 为带宽系数, p, p' 均为点的位置向量。

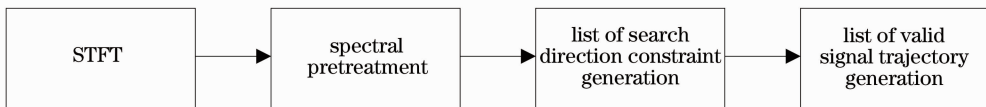


图 3 本算法的四个步骤

Fig. 3 Four steps of the algorithm

4 算法实现

根据本文第 3 部分的算法原理, 具体的算法实现流程可以分成如下四步执行。

4.1 STFT 频谱计算

根据本文第二节介绍的 STFT 的原理首先求

出原始 PDV 信号 STFT 时-频矩阵 $\mathbf{X}_k(\omega)$ ，一般而言， $\mathbf{X}_k(\omega)$ 的元素为复数，为了方便后续处理需要将 $\mathbf{X}_k(\omega)$ 转换为实数形式的功率强度值，转换公式如下：

$$\mathbf{P}_k(\omega) = 10 \lg[\mathbf{X}_k(\omega)\mathbf{X}_k^*(\omega)]. \quad (19)$$

由于在后面的运算过程中并不关心时间分量和频率分量的绝对大小，故用一般矩阵的表示形式来代替 $\mathbf{P}_k(\omega)$ ，即：

$$\mathbf{P}_{m \times n} = \mathbf{P}_k(\omega), \quad (20)$$

式中 m, n 分别表示时-频矩阵的行、列长度。

4.2 频谱预处理

从图 2 中可以看出，有效信号淹没在大量的噪声当中，不便于后期数据提取及处理。然而肉眼能从含有大量噪声的时频图像中识别出信号轨迹，是因为组成轨迹的点在坐标空间和强度空间具有一定的相似性，这种相似性可以作为轨迹搜索的指导原则之一^[10]，首先采用基于强度阈值的方法剔除大部分的功率谱低于一定阈值的频谱点，并将信号频谱图二值化，目的是提取出强度上相似性较高的频谱点，得到预处理过的时频矩阵 $\mathbf{D}_{m \times n}$ 作为后续算法的输入。具体操作步骤如下：

- 1) 统计原始信号的 STFT 频谱图 $\mathbf{P}_{m \times n}$ 的能量分布直方图 $H(I)$ ；
- 2) 计算 $H(I)$ 的累积分布直方图 $S(I)$ ；
- 3) 取满足条件 $S(I) \geq \eta S(I_{max})$ 的最小强度值 I 作为强度阈值 I_{Thr} ；
- 4) 如果 $\mathbf{P}_{m \times n}(i, j) \geq I_{Thr}$ ，那么 $\mathbf{D}_{m \times n}(i, j) = 1$ ，否则 $\mathbf{D}_{m \times n}(i, j) = 0$ 。

对于阈值 η 的选择要视信号的具体情况而定，一般而言，当信噪比低时， η 要选择的比较大，当信噪比高时， η 要选择的比较小，在本文中， η 设为 0.9。

4.3 生成搜索方向约束链表

设约束轨迹链表为 CList，其节点结构如下：

```
struct CNode
{
    Point p; // 连通域中心坐标
    intLabel; // 连通域编号
    CNode * next // 指针
};
```

为了从宏观上约束轨迹搜索沿着正确的方向进行，限定搜索方向的约束链表采用如下方式获得：首先用连通域提取算法^[11-13]，将整个时频图像内的所有连通的数据块 \mathbf{D}_i ，($i=1, 2, \dots, k$) 提取出来，并加以编号，设 \mathbf{D}_i 对应的编号为 i 。由于可能存在噪点，故限定每个连通区域内的数据个数不得低于某个阈

值 m_D ，若超过则舍弃该连通域数据，即连通域 \mathbf{D}_i 应满足条件：

$$m_i > m_D, i = 1, 2, \dots, k \quad (21)$$

对于点集 A ， $\exists a \in A$ ， $\exists A_1 \subset A$ ，定义：将元素 a 从 A 中删除的操作为 $A \ominus a$ ；将子集 A_1 从 A 中删除的操作为 $A \ominus A_1$ 。

约束轨迹搜索方向的算法实现流程如下：在 $\mathbf{D}_{m \times n}$ 上选择一个靠近轨迹起始点作为初始搜索点，记为 $p_s = (x_0, y_0)$ ，

1) $P_0 = p_s, T_0 = (0, 0)$ ，由 \mathbf{D}_C 的元素组成的约束轨迹链表 CList = NULL；

2) 搜索距离点 P_0 最近点 $P_1 = (xD'_1, yD'_1)$ ，如果 P_1 为空或者 $\text{Distance}(P_0, P_1) > D_{TC}$ ，则转到步骤 4)，否则计算 $T_1 = (xD'_1 - xD, yD'_1 - yD)$ ， \mathbf{D}_C 中删除点 P_1 ，即 $\mathbf{D}_C \ominus P_1$ ；

3) 若 $T_0 \cdot T_1 \geq \|T_0\| \|T_1\| \cos \zeta$ ，将 P_1 及其代表的连通域编号加入 CList 链表， $P_0 = P_1, T_0 = T_1$ ，否则跳到步骤 2)；

4) CList 生成完毕。

4.4 生成有效信号轨迹链表

设 GList 为信号曲线轨迹点链表，其节点结构为

```
struct GNode
{
    Point p; // 轨迹点坐标
    GNode * next // 指针
};
```

有效信号的轨迹链表生成算法如下：

1) $p_{MS1} = p_{MS2} = p_s, \text{GList} = \text{NULL}$ ；

2) 以 p_{MS1} 为中心，在 $\mathbf{D}_{m \times n}$ 中搜索其最近 m 邻域点集 D_m ，如果 D_m 为空，转到步骤 4)；否则以 p_{MS1} 为起点，以 D_m 为样本点集，进行 MeanShift 迭代，直至收敛于点 p_{MS} ， $p_{MS2} = p_{MS}$ ，收敛的条件为相邻两次迭代得到的点的距离小于常数 ϵ 。

3) 如果 $\text{Distance}(p_{MS2}, p_{MS1}) > D_{TC}$ ，转到步骤 4)；如果 $p_{MS2} = p_{MS1}$ ，说明前后两次搜索收敛于同一点，为了引导搜索方向远离 p_{MS2} ，以 p_{MS2} 为中心在 $\mathbf{D}_{m \times n}$ 点集中搜索其 k_M 最近邻域 \mathbf{D}_{km} ， $\mathbf{D}_{m \times n} \ominus \mathbf{D}_{km}$ ，转到步骤 2)；否则 $p_{MS1} = p_{MS2}$ ，将 p_{MS2} 加入 GList 链表；

4) GList 生成完毕。

事实上，4.2 节实现的是信号强度空间上的相似性搜索，4.3 和 4.4 节实现的是信号坐标空间上的相似性搜索，整个算法结构如图 4 所示。

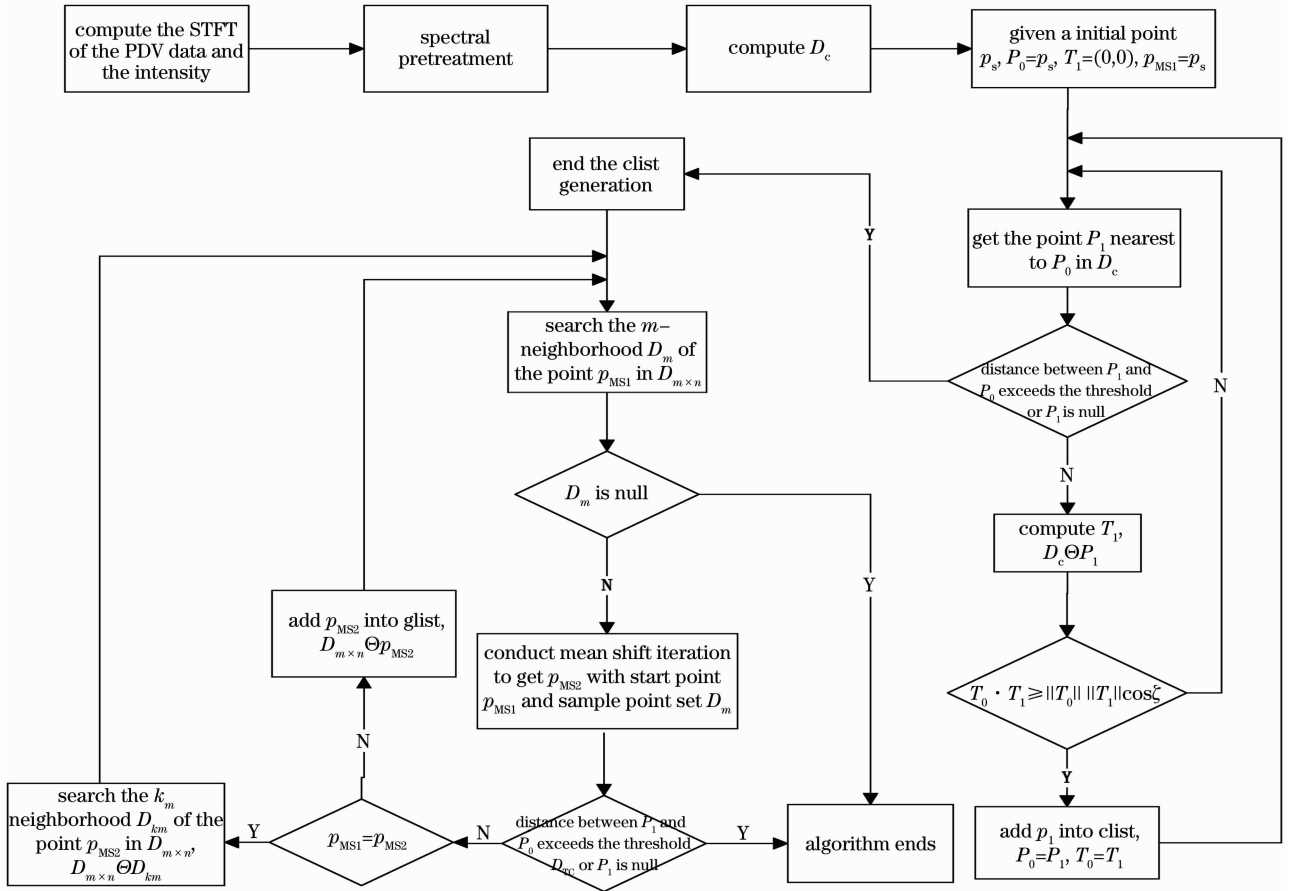


图 4 整体算法结构图

Fig. 4 Overall structure of the algorithm

5 实验结果及分析

图 5 和图 9 所示的两组实验数据来源于 Risi RP-80 雷管引爆瞬间的冲击波的激光多普勒差频信号,实验系统采取了激光波长为 $1.55 \mu\text{m}$ 的全光纤结构^[14-16],实验数据处理平台为 Matlab R2012b,计算机采用 AMD 双核 2.1 GHz CPU,2G 内存的配置。由于实验采用了 50 GHz 的示波器对数据进行

采样,在非常短的爆炸时间内可以获得足够多的样本数据,本次实验的样本长度 2000000 个,窗函数为海明窗,窗口长度为 256,每段重叠样本数为 200。图 8 和图 13 所示为采用本文算法提取的有效信号轨迹曲线。图 9 和图 14 所示为采用传统的 STFT 频谱极大值方法提取的信号轨迹。

从以上两组实验结果可以看到:采用传统的基

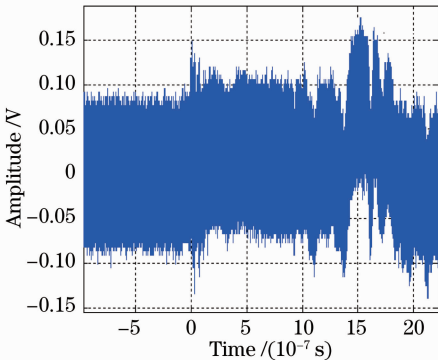


图 5 第一组爆炸冲击波的 PDV 数据

Fig. 5 First group of the PDV data of shock wave

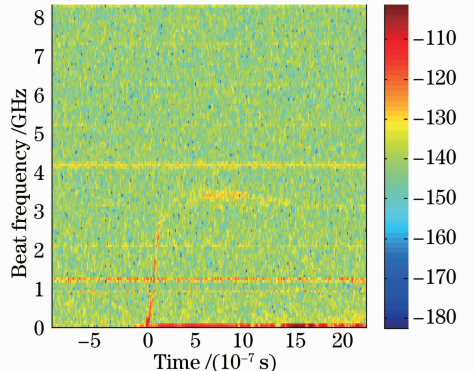


图 6 第一组数据的 STFT 频谱图

Fig. 6 STFT spectrum of the first group of the PDV data

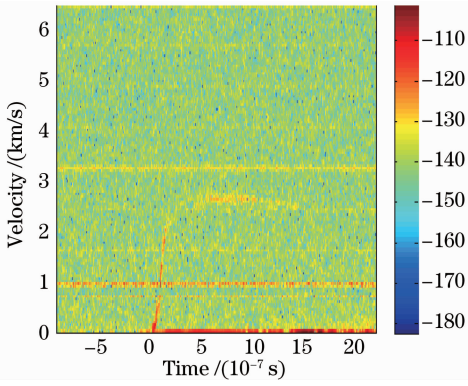


图 7 由图 6 换算出的速度-时间图

Fig. 7 Velocity-time diagram translated from Fig. 6

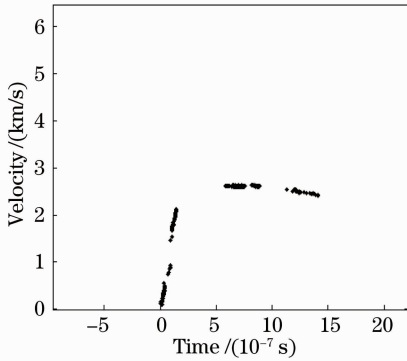


图 8 采用本算法对图 7 进行轨迹搜索的结果

Fig. 8 Result when conduct trajectory search with the diagram shown in Fig. 7

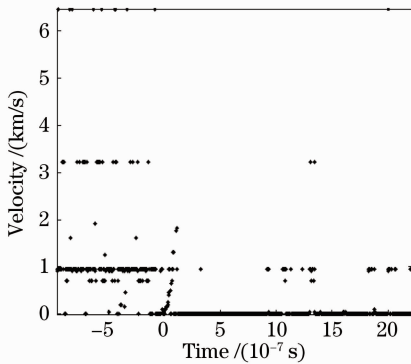


图 9 采取 STFT 频谱极大值方法对图 7 进行信号提取的结果

Fig. 9 Result when conduct the method of STFT spectral maxima with the diagram shown in Fig. 7

于 STFT 频谱极大值的方法提取的信号绝大部分为背景干扰和噪声,几乎不含有有效的速度信号(见图 9 和图 14),而采用本文所提出的基于 MeanShift 的信号轨迹搜索算法提取的信号却基本不含有噪声,并且比较好地保留了有效信号的完整性(见图 8 和图 13)。在图 8 和图 13 中,相比于原始的信号频谱图(见图 7 和图 12),相应的信号轨迹

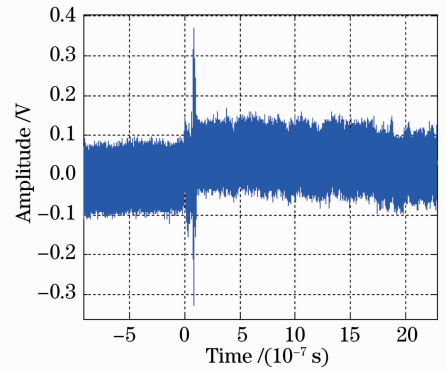


图 10 第二组爆炸冲击波的 PDV 数据

Fig. 10 Second group of the PDV data of shock wave

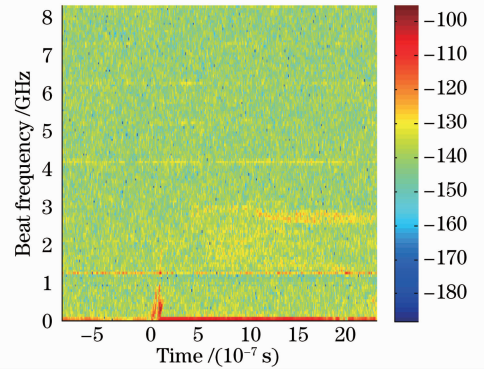


图 11 第二组数据的 STFT 频谱图

Fig. 11 STFT spectrum of the second group of the PDV data

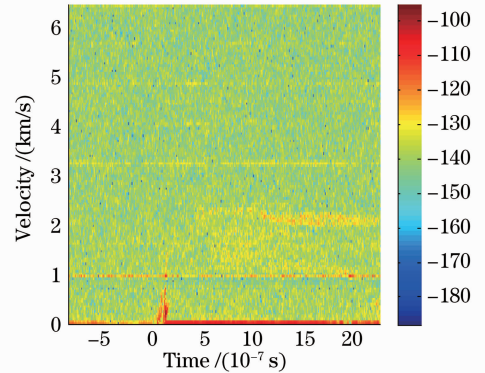


图 12 由图 11 换算出的速度-时间图

Fig. 12 Velocity-time diagram translated from Fig. 11

曲线在顶部位置有部分断裂,产生这种现象的主要原因有由于断裂部分的频谱信号强度比较低(颜色较浅),在经过阈值滤波后较低强度的频谱点被直接滤除掉了,事实上,阈值滤波步骤中阈值 η 的选取对信号完整性和信号信噪比具有直接影响,当阈值选得过低,有效信号轨迹的完整性虽然得到保证,但是 $D_{m \times n}$ 中将含有大量噪声,有可能使得搜索方向约束链表偏离正确的方向,导致最终的轨迹搜索结果含有噪声点,甚至得不到正确的轨迹曲线,也即信号的

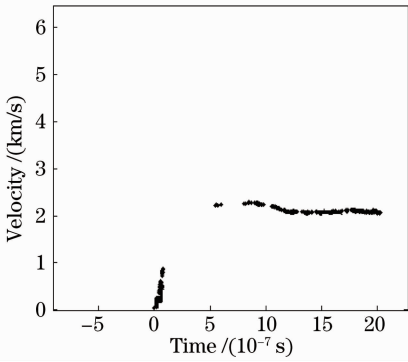


图 13 采用本算法对图 12 进行轨迹搜索的结果

Fig. 13 Result when conduct trajectory search with the diagram shown in Fig. 12

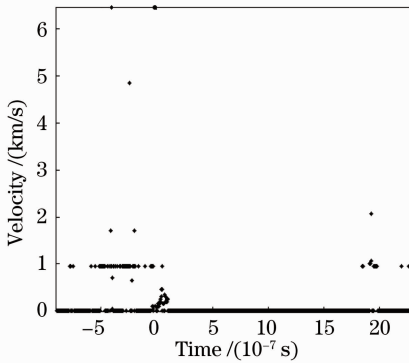


图 14 采取 STFT 频谱极大值方法对图 12 进行信号提取的结果

Fig. 14 Result when conduct the method of STFT spectral maxima with the diagram shown in Fig. 12

信噪比降低;当阈值选的过高,虽然大部分噪声得到有效滤除,但是也将导致有效信号的部分轨迹也被滤除,使得信号的完整性受损。在实际应用中应根据原始信号的具体情况选择合适的阈值。

在实际测量中得到的多普勒频率并不是单一的频率,而是具有一定的宽度,这种现象被称为多普勒信号加宽^[17]。在图 8 和图 13 中确实可以看到信号轨迹在其法线方向有一定的弥散现象,如何减小多普勒加宽效应并非本文内容,故在轨迹搜索结果中,仍然保留了信号加宽这一特性。

从时频图可以看出,有效信号淹没在大量的噪声当中,而且还有干扰直线穿过有效信号的轨迹曲线,但是采用本文所提出的基于 MeanShift 的轨迹搜索算法所提取的信号轨迹却几乎不含有噪点,并且轨迹也没有转到错误的方向去。原始信号轨迹虽然出现间断现象^[18],但是信号轨迹还是能够比较完整地提取出来,较好地保留了原信号的特征(见图 8 和图 13),而采用传统的基于 STFT 频谱极大

值的方法提取得到的信号的绝大部分为背景干扰和噪声,几乎不含有有效的速度信号(见图 9 和图 14),可见采用 STFT 频谱极大值提取算法处理本文中给出的爆炸冲击波的多普勒信号效果很不理想,这是由于爆炸环境复杂多变,测量过程中不可避免地受到大量杂散碎片微粒以及背景的干扰,如果这些噪声干扰强度超过了待测目标的信号强度,那么基于频谱幅值大小的 STFT 频谱极大值提取算法就会失效,这是因为这种方法默认噪声干扰的强度不会超过有效信号的强度,事实上,所有仅基于信号强度大小的提取算法在处理这一类的信号时均会失效。本文提出的基于 MeanShift 的轨迹搜索算法首先采用一个较低的阈值滤除频谱中的一部分噪声(但不至于将有效信号滤除),然后引入搜索方向约束链表来引导轨迹搜索沿着正确的方向进行,由此可见,本文提出的算法不仅仅基于信号强度的判断,而且考虑了信号轨迹的空间几何形态,因此有较强的抗噪性能,比传统的 PDV 信号提取方法有更加广泛的适用范围。

从表 1 可以看出,在算法耗时方面,基于 MeanShift 的有效信号提取算法的效率比传统的 STFT 频谱极大值的提取方法要低,这是因为基于 MeanShift 的有效信号提取算法需要执行 STFT 频谱计算、频谱预处理、生成搜索方向约束链表、生成有效信号轨迹链表四个步骤,STFT 频谱计算和频谱预处理两个阶段的算法的执行流程与 STFT 频谱极大值方法基本一致,所以两者的耗时量比较接近。而生成搜索方向约束链表和生成有效信号轨迹链表这两个步骤中含有大量的最邻近点搜索操作以及 MeanShift 迭代操作,这是本算法的主要耗费时间的环节。从表 2 中可以看到,采用本算法处理两组实验数据在这两个步骤在上消耗的时间分别占到了总时间的 77% 和 87%。

从图 6 和图 11 中可以看出,第二组数据的噪声干扰比第一组数据的要严重,相应的算法执行时间也比第一组的长,一般而言,数据的强噪声干扰越多,算法的耗时也越长,这是因为强噪声干扰的增加导致了 $D_{m \times n}$ 和 D_i 数目的增加,扩大了算法进行轨迹搜索的样本空间容量。从表 1 中可以看出,无论实验数据的噪声水平如何,STFT 频谱极大值方法所消耗的时间基本不变,这是由于该方法直接提取 STFT 频谱中每一个时间点上的频率极大值点作为有效信号点,噪声的多少基本上并不会影响到算法执行时间的长短。

表 1 两种方法耗时情况对比

Table 1 Contrast to time consuming situation for these two methods

Methods of PDV signal processing	Processing time in the first set data /ms	Processing time in the second set data /ms
Method of MeanShift	1478.8	2678.5
Method of STFT spectral maxima	403.2	415.6

表 2 MeanShift 信号提取方法四个阶段的耗时情况

Table 2 Situation of time consuming in the four stages of the MeanShift signal extraction method

Time consuming	STFT spectral	Spectral	Direction constraint	Signal trajectory list
	computation /ms	pretreatment /ms	list generation /ms	generation /ms
First set data	201.3	33.5	409.0	735.0
Second set data	209.2	37.9	520.4	1811.0

6 结 论

讨论了目前传统的 PDV 信号提取算法在抗噪性能上不足的原因,并且提出了一种基于 MeanShift 的 PDV 信号提取算法,阐述了相关的原理以及具体的算法实现。分别采用本文提出的算法与传统的 STFT 频谱极大值信号提取算法对两组有强噪声干扰的 PDV 实验数据进行处理,实验结果表明,传统的 STFT 频谱极大值信号提取算法在强噪声条件下会失效,而本文提出的算法却能在大量的强噪声干扰中将有效信号提取出来,体现了较好的抗噪性能。在算法执行效率上,本文提出的算法比 STFT 频谱极大值方法要低,并且,本算法的执行时间与数据的噪声水平呈正相关的关系,噪声干扰越严重,算法的执行时间越长,但依然能够准确地从 STFT 频谱中提取出有效信号的轨迹,在处理含有强噪声干扰的 PDV 信号方面较之传统的 PDV 信号处理算法具有更广泛的适用范围。

参 考 文 献

- 1 D H Dolan. Accuracy and precision in photonic Doppler velocimetry [J]. Review of Scientific Instruments, 2010, 81(5): 053905.
- 2 Du Jun, Qu Yancheng, Zhao Weijiang, *et al.*. Measuring error of phase modulation laser Doppler shift measuring method[J]. Chinese J Lasers, 2014, 34(7): 0712001.
杜 军, 曲彦臣, 赵卫疆, 等. 相位调制激光多普勒频移测量方法的误差研究[J]. 中国激光, 2014, 34(7): 0712001.
- 3 O T Strand, D R Goosman, C Martinez, *et al.*. Compact system for high-speed velocimetry using heterodyne techniques [J]. Review of Scientific Instruments, 2006, 77(8): 083108.
- 4 Wang Yuegang, Ji Shao, Xu Hongtao. Non-stationary signals processing based on STFT [C]. The Eighth International Conference on Electronic Measurement and Instruments, 2007, 3(1): 301-304.
- 5 Michael Furlanetto. Velocity extraction from PDV data using advanced Fourier transform techniques [C]. APS Shock Compression of Condensed Matter Meeting Abstracts, 2009, 1(1): 3003.
- 6 Shouxian Liu, Detian Wang, Tao Li, *et al.*. Analysis of photonic Doppler velocimetry data based on the continuous wavelet transform

- [J]. Review of Scientific Instruments, 2011, 82(1): 023103.
- 7 Songhua Wu, Zhishen Liu, Dapeng Sun. Noise reduction in LOS wind velocity of Doppler lidar using discrete wavelet analysis[J]. Chin Opt Lett, 2003, 1(12): 722-725.
- 8 Yizong Cheng. Mean shift, mode, seeking, and clustering[C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.
- 9 Dorin Comanicin, Peter Meer. Mean shift analysis and applications[C]. IEEE Transactions on Computer Vision, 1999, 2(1): 1197-1203.
- 10 S Z Selim, M A Ismail. K-means-type algorithms: a generalized convergence theorem and characterization of local optimality[C]. IEEE Transaction on Putter Analysis and Muchine Intelligence, 1984, 6(1): 81-87.
- 11 Chen Baisheng. A new kind of method on marking the connected region of binary image [J]. Computer Engineering and Applications, 2006, 25(10): 46-47.
陈柏生. 一种二值图像连通区域标记的新方法[J]. 计算机工程与应用, 2006, 25(10): 46-47.
- 12 Kenji Suzuki, Isao Horiba, Noboru Sugie. Linear-time connected-component labeling based on sequential local operations[J]. Computer Vision and Image Understanding, 2003, 1(89): 1-23.
- 13 Prabir Bhattacharya. Connected component labeling for binary images on a reconfigurable mesh architectures [J]. J Systems Architecture, 1996, 4(42): 309-313.
- 14 Jian Wang, Chongqing Wu, Hongwei Song, *et al.*. Measurement of the velocities in the transient acceleration process using all-fiber photonic Doppler velocimetry[C]. SPIE, 2011, 7753: 77536F.
- 15 Sui Xiaolin, Zhou Shouhuan, Zhao Hong, *et al.*. Research on all-fiber structure velocity measurement coherent lidar[J]. Chinese J Lasers, 2013, 40(12): 1208007.
睦晓林, 周寿桓, 赵 鸿, 等. 一种全光纤结构的相干激光测速雷达研究[J]. 中国激光, 2013, 40(12): 1208007.
- 16 Wang Xitao, Liu Bingyi, Wu Songhua, *et al.*. 1.55 μm all-fiber laser heterodyne detection and data analysis with high measurement accuracy of velocity[J]. Laser & Optoelectronics Progress, 2011, 48(6): 060301.
王希涛, 刘秉义, 吴松华, 等. 高精度 1.55 μm 全光纤激光相干测速实验及数据分析[J]. 激光和光电子学进展, 2011, 48(6): 060301.
- 17 Zhou Jian, Yao Baoju, Long Xingwu. The study on the transit widening of laser Doppler signal [J]. Infrared and Laser Engineering, 2011, 40(5): 826-829.
周 健, 姚宝聚, 龙兴武. 激光多普勒信号渡越加宽研究[J]. 红外与激光工程, 2011, 40(5): 826-829.
- 18 Zhu Jianzhong. The tracing algorithm of the intermittent curves [J]. Computer Applications, 1994, 4(1): 27-28.
祝建中. 间断曲线的追踪算法[J]. 计算机应用, 1994, 4(1): 27-28.