

全波形机载激光数据分解研究^①

刘 赞^{②*} 张爱武^{③*} 段乙好^{*} 王书民^{**}

(^{*}首都师范大学三维信息获取与应用教育部重点实验室 北京 100048)

(^{**}中国地震局地震预测研究所 北京 100036)

摘要 针对进行全波形机载激光数据分解时采用重力中心(COG)方法和高斯脉冲拟合(GPF)方法难以检测到全波形数据中的叠加波和弱波脉冲,从而造成地物信息丢失的问题,提出了一种改进的严格高斯检测(RGD)算法。该算法首先采用高斯滤波提高信噪比,然后通过脉冲检测估计初始参数,并利用 Trust Region 参数优化方法进行高斯拟合,最终有效地分解出波形中的叠加波和弱波。利用提取到的每个单次回波的标准差、振幅、位置等模型参数能够得到半高波宽(FWHM)、后向散射截面和回波强度等波形信息。多次试验表明,与 RGD 算法相比,该方法可以有效地提高叠加波的识别率,减小波形拟合误差。

关键词 全波形,叠加波和弱波,波形分解,RGD 算法,Trust Region

0 引言

全波形激光雷达所记录的全回波信息是发射激光脉冲沿途遇到的所有目标回波信号的总和,反映了地物垂直分布,揭示了地物的几何和物理属性,是地物分类的重要判据。如何从全波形数据中准确分解出单次回波,并提取振幅、距离、半高波宽(即最大振幅 1/2 处对应的全宽,英文为 full width at half maximum,缩写为 FWHM)、后向散射截面等波形信息则成为近期全波形激光雷达数据处理的热点问题。Zhu^[1]提出了逐步波形分解(progressive waveform decomposition, PWD)方法,该方法对单次回波采取逐一提取的策略,难以处理只有一个主要峰值点存在的叠加波形,且使用拐点进行初始值估计,鲁棒性低;Ma^[2]采用改进的最大期望(expectation maximum, EM)算法进行参数优化,与传统的非线性优化算法相比,它不需要最小二乘拟合过程,但使用一阶导数进行波形的检测,难以识别不太明显的叠加波;Hofton^[3]引入“重要性”的概念,对每个高斯成分进行标记和排序,根据高斯成分个数最小化的原

则进行优化,但面对复杂的叠加波形,难以有效而准确地寻找拐点。Chauve^[4]考虑到非对称的波形,引入广义高斯模型,但大多数情况脉冲波形“对称”的假设是有效的,因此采用标准的高斯函数更为保守。Roncat^[5]利用均方根误差函数(ASDF)来检测回波,能够抑制因受激振(ringing)效应影响而检测到假回波,但该方法对脉冲波形并不敏感;Roncat^[6]引入 B 样条函数来拟合脉冲波形,通过反卷积得到每个单次回波对应的不同目标的散射截面。该方法与高斯分解方法相比,既不需要初值也不依赖于对称的波形,但复杂度增加,且只能获取不同目标的反射截面。Still^[7]通过建立波形中的局部临近关系来检测弱波,但未考虑叠加波;Yu^[8]提出了一种严格脉冲检测(rigorous Gaussian detection, RGD)算法,相比传统的检测方法,该方法大大提高了弱波与叠加波的识别率。但是,直接对观测波形进行处理会不可避免地受到噪声的影响,同时,面对复杂的脉冲叠加情形,该方法也显得不足。本研究针对上述问题进行了探索,在此基础上提出了改进型的 RGD 算法,进一步提高了叠加波的识别率。

① 国家科技支撑计划(2012BAH31B01)和北京市自然科学基金重点项目(B类)(KZ201310028035)资助项目。

② 男,1986 年生,硕士;研究方向:机载激光雷达全波形数据的分解;E-mail:liuzh07@163.com

③ 通讯作者,E-mail:zhangaw98@163.com

(收稿日期:2013-07-23)

1 改进的 RGD 算法

RGD 算法,也称为严格的脉冲检测算法,它基于常用的高斯分解方法,但通过严格的初始值估计和复杂的迭代过程来实现。算法通过分析波形的二次导数和明显峰值点两侧的样本个数来寻找叠加波,利用高斯拟合的高残差值来寻找漏掉的弱波。

本文在此方法基础上提出了改进的 RGD 算法,它的不同之处在于增加了高斯滤波操作以降低随机噪声的影响。同时,提出通过寻找负的二次导数的方法来检测叠加波。算法的优势在于可以更有效地检测只有一个主要峰值点存在的叠加波形,且检测时不受回波脉冲叠加个数限制。改进算法的分解流程如图 1 所示,分为高斯滤波、弱波与叠加波的检测、波形拟合与分形分解、粗差剔除等过程。

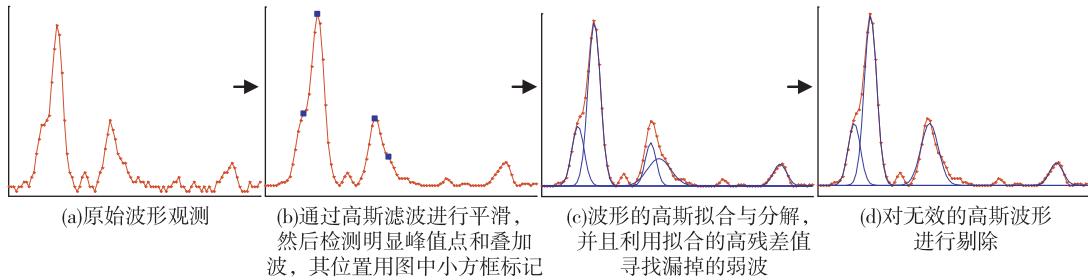


图 1 改进的 RGD 算法的流程

1.1 高斯模型与高斯滤波

根据 Wagner^[9]的理论,一个回波信号通常通过高斯模型来表示。而接收信号是发射激光脉冲沿途遇到的所有目标回波信号的总和,由多个目标回波叠加而成,考虑到背景噪声,可用下式表达:

$$\varphi_i(x) = A_i \cdot \exp\left[-\frac{(x - \mu_i)^2}{2\delta_i^2}\right] \quad (1)$$

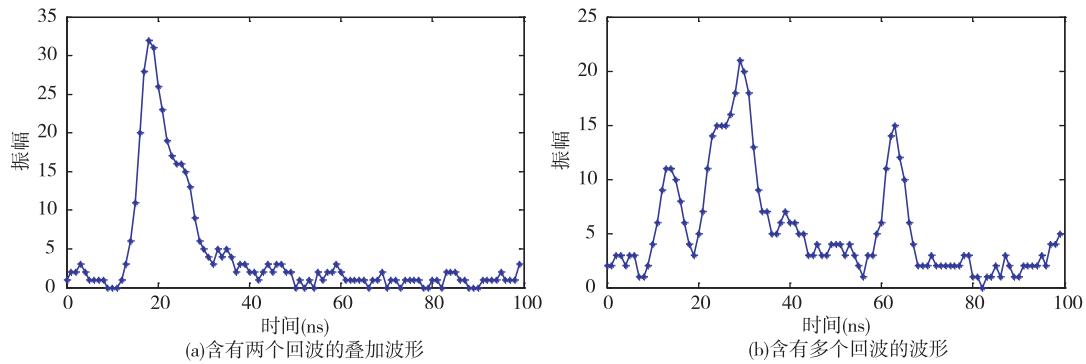
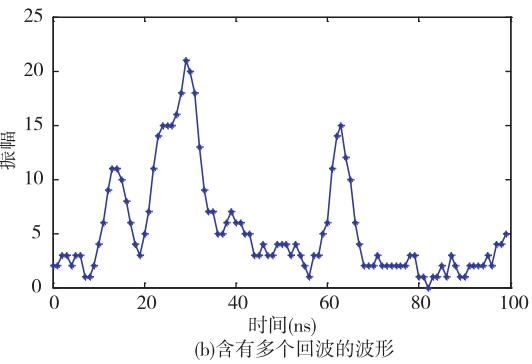


图 2 原始波形

从图 2 可以看到,观测波形中有多个回波,它记录了地物的垂直分布,因此,需经过波形分解过程才能分解出单次回波,提取振幅、距离、半高波宽(FWHM)、后向散射截面等波形信息,从而了解地物的垂直结构特征。因原始波形中存在背景噪声和随机噪声,本文采用一维高斯滤波方法进行降噪处理。滤波函数选用高斯函数作为模板,模板宽度选

$$P_r(x) = \sum_{i=1}^N \varphi_i(x) + b \quad (2)$$

其中 $\varphi_i(x)$ 为第 i 个高斯函数分量, A_i 、 μ_i 、 δ_i 分别代表该高斯函数的权重、均值和标准差, N 为高斯函数个数, b 为背景噪声, $P_r(x)$ 为 N 个高斯函数的加和。原始波形如图 2 所示。



择为 2ns(发射脉冲的半宽),窗口为 1×3 。滤波之前对原始波形数据进行预处理,小于分离噪声的阈值 A_n 的所有采样点的值都重新给定为 A_n ,其中 $A_n = 3\delta_n$, δ_n 为噪声标准差,通过统计将近 5000 条回波脉冲波形记录中每条记录最后 10 个样本点的值来计算标准差。

1.2 弱波与叠加波的检测

在回波脉冲检测上,常用的方法有^[10-12]:(1)峰值检测;(2)边缘检测;(3)恒比检测;(4)重力中心法检测。这些都是基于阈值的。本文首先采用峰值检测来寻找具有明显峰值点的回波脉冲。基于 Stilla^[7]对发现新的脉冲波形的定义,如果一个波形在连续至少 5ns 的间隔内其振幅值都大于 3 倍的噪声标准差,即 A_n ,则被接受。原 RGD 算法利用二阶导数的极值点来寻找不太明显的叠加波。本文对叠加波形进行模拟分析,模拟过程如图 3 所示,从图 3(b)

可看出当两个回波脉冲距离为 4ns 时,只存在一个主要的峰值点;而当距离等于 3ns 时,如图 3(c)中所示,只存在一个脉冲波形。根据王燕^[13]定义的离散点二阶导数的五点数值微分公式,计算样本点的二次导数,发现只有一个明显峰值点存在的叠加波形(如图 3(b)中所示)叠加波对应位置处均为二阶导数的极小值点,并且值小于零。因此明显峰值点两端负的二阶导数极小值点表明了叠加波的存在。

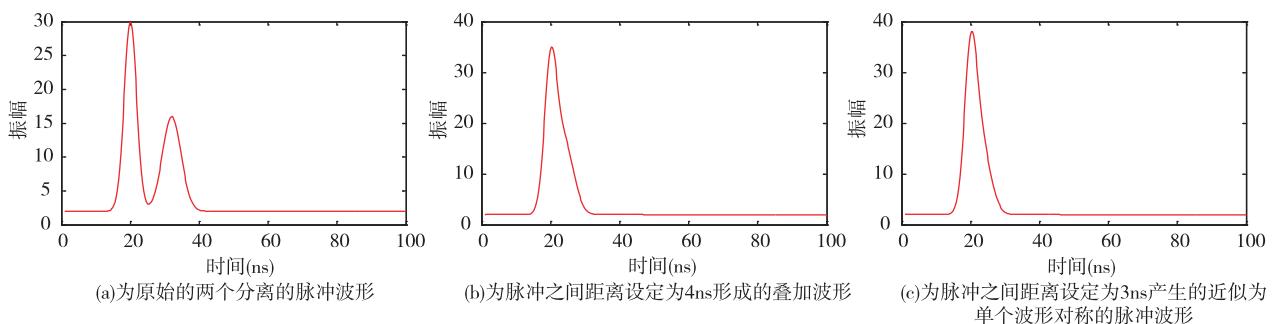


图 3 两个不同距离、不同振幅的脉冲所模拟的接收波形

(1) 明显峰值点的检测

通过局部极大值法来寻找回波脉冲的峰值点,峰值点可分为弱波峰值点和非弱波的峰值点。需满足的条件为: $X_i > A_n$ 且 $X_{i-1} < X_i \geq X_{i+1}$ 且 $X_{i-2} < X_{i-1}$ 且 $X_{i+2} \leq X_{i+1}$, 其中 X_i 代表第 i 个样本点,并且要求 5 点的样本值均大于 A_n 。进一步判断:若 $X_i \geq A_i$ (其中 A_i 扫描系统的检测阈值, 小于此值的脉冲被认为是弱波脉冲), 则 X_i 这一样本点为潜在的明显峰值点;若 $X_i < A_i$, 则此点可能为弱波的明显峰值点,为了避免错误的检测,进一步要求 $X_{i-4} < X_{i-2}$ 或者 $X_{i+4} < X_{i+2}$ 。完成以上步骤后,进入最终的判断:若检测到的潜在峰值点与已有的峰值点之间间隔大于等于 4ns,则新的回波脉冲被发现。

(2) 叠加波的检测

完成明显峰值点的检测后,通过寻找二阶导数值为极小值且小于 0 的点来检测叠加波,检测的流程如图 4 所示。检测前对明显峰值点左右两侧大于 A_n 的样本点个数分别进行统计,若某一侧的个数大于等于 5,则对其进行检测。步骤(a)中利用了 Yu^[8]提出的假设:检测到的峰值点的振幅值应大于 1/3 明显峰值点的振幅值。步骤(b)判断可能存在的一些不明显叠加波。步骤(c)和步骤(d)中的条件设置避免了错误的检测,若满足步骤(d)中的条件,表

明检测到叠加波,给定其位置位于波形中有叠加波的那一侧的末端点与上个峰值点之间的中间位置处。

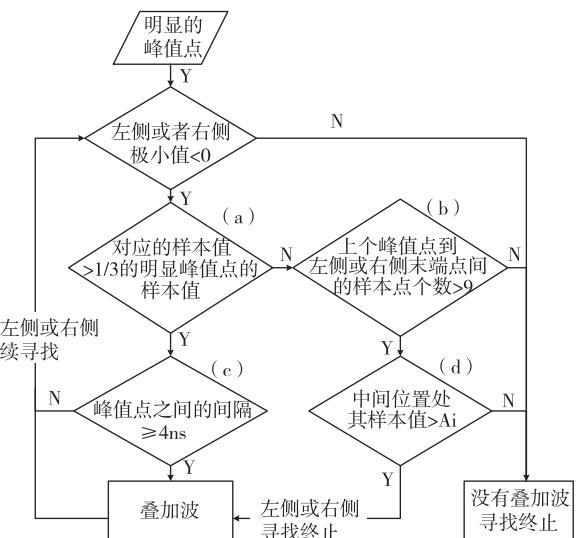


图 4 寻找叠加波的流程

1.3 波形拟合与分形分解

完成峰值点的检测后可判断检测到的脉冲个数,将峰值点的振幅和所处位置作为相应回波脉冲的初值,脉冲的波宽初值给定为发射脉冲的波宽值,

从而完成初始值的估计,之后利用非线性优化算法进行波形的拟合与参数的迭代优化。并在拟合过程中寻找漏掉的弱波。典型的优化算法有二种:(1)利用 LM 优化算法的非线性最小二乘(NLS)^[3];(2)利用最大期望(EM)算法的最大似然估计^[14]。由于 LM 算法在迭代优化时会出现因遇到局部最小值而无法找到最优解的情况,因此本文选择 Nocedal^[15]定义的基本的 Trust Region 方法,并采取一定的策略来计算增长步。定义目标函数为 F ,我们假定一个模型 L ,目标函数在当前的迭代 x 处的泰勒展开式为

$$F(x + h) \approx L(h) = F(x) + h^T c + \frac{1}{2} h^T B h \quad (3)$$

其中 x 为变量的向量, h 为 x 的增量, $c \in \mathbf{R}^n$, $c = F'(x)$, $B = F''(x)$, $B \in \mathbf{R}^{n \times n}$, B 为对称矩阵。优化算法关键步骤是寻找迭代参数的增长步 h 。本文采取这样的策略:当 B 为正定且 $\|B^{-1}g\| \leq \Delta$ 时,其中 Δ 为信任区域半径,则 $h = -B^{-1}g$ 为增长步;若不满足则采用 More^[16]提出的计算增长步的算法 3.14。完成参数的优化后,根据波形拟合后与原始波形之间的高残差值来寻找漏掉的回波脉冲,若残差大于等于 A_n ,且此点不是因激振效应产生的,则在此位置增加一个新的回波脉冲。同时,在高斯拟合时为了避免脉冲波形以外的样本点进入最小二乘运算,只取每个高斯成分峰值处的两侧对应的样本点进入最小二乘逼近。

1.4 粗差剔除

优化完成后需要通过设置条件删掉无效的高斯成分。原 RGD 算法中规定两个脉冲波形之间距离的阈值为 2ns,根据图 3 中的模拟实验,同时考虑到本文算法加入了滤波操作,发现取阈值为 3ns 更为合适。条件有:(1) $Sep < 3\text{ns}$; (2) $A < A_n$; (3) $\delta < 0.7\delta_0$; (4) $A < 1/13\max(\mathbf{x})$ 且 $\max(\mathbf{x}) > Ar$ 且 $Ds > 0$; (5) $A < A_i$ 且 $\delta > 1.5\delta_0$ 且 $\delta < 0.9\delta_0$ 。其中 Sep 代表两个高斯脉冲间的距离, δ 代表回波脉冲高斯模型的标准差; δ_0 代表发射脉冲高斯模型的标准差; A 代表高斯脉冲峰值点的振幅; $\max(\mathbf{x})$ 为波形中样本点最大的振幅值; Ar 为判断是否是激振效应的阈值; Ds 代表高斯脉冲峰值点处与波形中样本点最大的振幅值对应的位置之间的差; A_i 系统的检测阈值。只要满足这 5 个条件中之一,则判断为无效的脉冲,删除。因算法拟合的是滤波后的波形,因此,最后需对原始的观测波形通过高斯拟合进行逼

近,从而完成全波形数据的分解。

2 结果与分析

2.1 典型波形检测结果的对比

本文实验的激光雷达全波形数据由 Riegl LMS Q-560 提供,采样间隔为 1ns,采样数为 100。为了检验叠加波形检测结果,选取了 8 个典型的波形记录进行分析,图 5 为通过本文算法检测得到的结果。

分别和商业软件的重力中心(center of gravity, COG) 法和高斯脉冲拟合(Gaussian pulse fitting, GPF) 法以及 RGD 算法作对比,检测结果如表 1 所示。可看到 COG 方法和 GPF 方法的检测除第 7、8 条记录外,其他波形记录中的叠加波都被漏掉。由此可见这两个算法本身还存在缺陷;RGD 算法检测到多数叠加波,但是第 3、4 条波形记录中的叠加波被漏掉,第 5 条记录中波形右侧有 3 个叠加的脉冲,只检测到 2 个,表明对于两个以上脉冲在同一侧叠加的情形,RGD 算法还难以处理;而本文改进的检测算法对不同的脉冲叠加情形均体现出了有效性。

2.2 统计分析

分别选取了 150 个叠加波形的回波记录和有弱波存在的回波记录进行检测,其叠加波的形状如图 5 中的 8 个典型波形,之后进行对比分析。

2.2.1 叠加波的检测结果

表 2 给出了不同算法统计结果的对比,相比前两个方法,RGD 算法与本文算法都能够有效地检测出叠加波脉冲,但 RGD 漏掉了接近 1/3 的叠加波,而本文算法检测到所有的叠加波。这验证了针对不同的脉冲叠加情形改进算法的有效性及优势。

RGD 算法未检测到的叠加波主要可分三类,分别如图 6 中的(a)、(b)、(c) 所示,其中(a)为叠加的弱波,(b)为不太明显的叠加波,(c)为两个回波脉冲在明显峰值点某一侧叠加。可以看到图 6(a) 中叠加的弱波被漏掉;图 6(b) 中的叠加波被漏掉;图 6(c) 中在波形右侧有两个叠加波因算法本身的限制只提取出一个。

为了对比 Trust Region 和 LM 优化算法所拟合结果的精度,对 150 个叠加波形的回波记录使用这两个算法分别进行拟合,并且采用相同的迭代收敛条件,计算拟合结果的均方根误差(RMSE)。本文选取大于 Dc 补偿值 2 的样本进入均方根误差的计算,尽量避免背景噪声对计算的影响。两者拟合得到的 RMSE 的差值如图 7 所示。

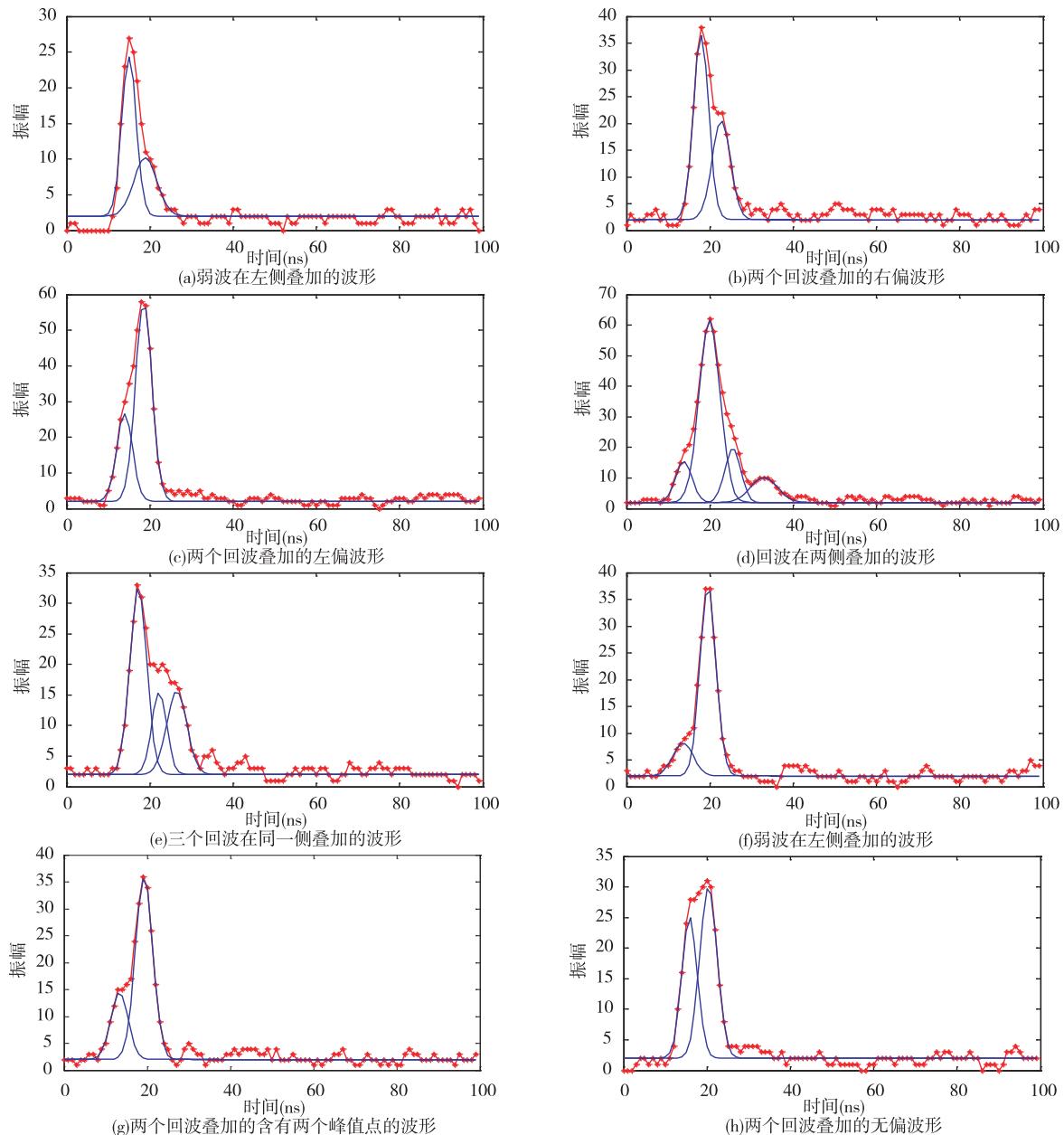


图 5 回波脉冲检测结果

表 1 不同脉冲检测方法结果对比

波形记录 编号	COG 算法 检测个数	GPF 算法 检测个数	RGD 算法 检测个数	改进算法 检测个数
1	1	1	2	2
2	1	1	2	2
3	1	1	1	2
4	2	2	3	4
5	1	2	2	3
6	1	1	2	2
7	1	2	2	2
8	1	2	2	2

表 2 不同算法统计结果对比

脉冲检测 方法	从叠加波 中检测到 0 个回波 脉冲的 个数	从叠加波 中检测到 1 个回波 脉冲的 个数	从叠加波 中检测到 2 个回波 脉冲的 个数	从叠加波 中检测到 3 个回波 脉冲的 个数	检测率 (%)
COG 算法	2	142	6	0	4
GPF 算法	0	123	27	0	18
RGD 算法	0	45	104	1	70
本文算法	0	0	146	4	100

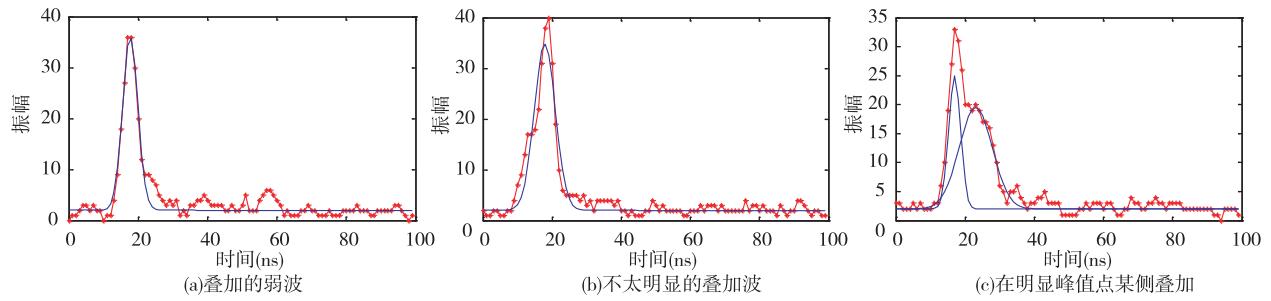


图 6 RGD 算法未能检测的三种脉冲叠加情形

均方根误差用式

$$RMSE = \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / N} \quad (4)$$

表示,其中 N 为样本的个数, y_i 为第 i 个观测值, \hat{y}_i 为第 i 个拟合值, $RMSE$ 为均方根误差。

可以看出 Trust Region 和 LM 优化方法两者均方根误差的差值都是在 10^{-9} 的数量级上,对于波形分解结果而言,这样的差异完全可以忽略不计。相比小于 0 的部分,大于 0 的部分更多一些,这说明 Trust Region 拟合得到的均方根误差更大一些,但 Wagner^[9]指出使用 LM 算法优化可能会产生负的振幅,并且在某些情况下难以找到解,而 Trust Region 总能寻找到合适的解,因此采用 Trust Region 方法优化结果更加稳健。总的来说,Trust Region 能避免错误或不准确的分解结果,降低波形分解的拟合误差。

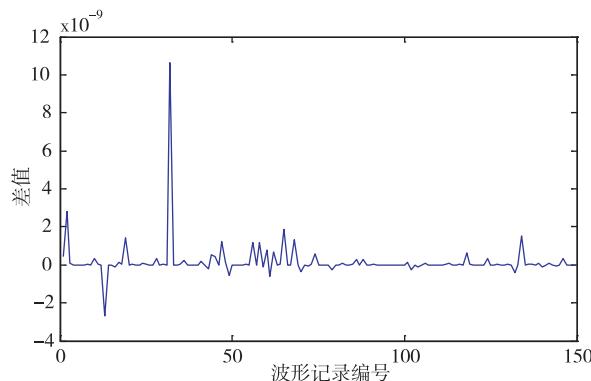


图 7 优化算法所拟合结果的 RMSE 差值

2.2.2 弱波的检测结果

表 3 给出了不同算法对弱波的检测结果对比,从中可看出 COG 与 GPF 方法都只检测到少数的弱波,都是振幅处于阈值水平的弱波。而 RGD 算法和本文算法检测到大部分弱波,因为在弱波的检测上其检测算法基本相同,所以结果相差并不大。

表 3 不同算法对弱波的检测结果对比

脉冲检测方法	未检测到弱波的个数	检测到弱波的个数	检测率(%)
COG 算法	90	60	40
GPF 算法	78	72	48
RGD 算法	21	129	86
本文算法	16	134	89.33

2.3 仿真验证

实测数据没有真实的高斯函数值作为比对,因而采用模拟信号对叠加波和弱波的检测结果进行仿真验证,验证过程如图 8 所示。受噪声的影响,从图 8 的(d)图中看出弱波的分解结果与模拟的原始高斯函数有一定差异,但总体上能够很好地近似复原,结合文中对大量实测数据分解结果的检查,验证了改进的 RGD 算法在叠加波与弱波检测上的有效性。

通过以上的统计分析,改进的 RGD 算法能够有效地检测叠加波和弱波脉冲,提高叠加波和弱波识别率。波形分解后能够得到每个地物回波脉冲的振幅、位置和标准差等模型参数,进而计算得到半高波宽($FWHM$)、后向散射截面和回波强度等波形信息。

这些波形信息通过以下公式得到。半高波宽 $FWHM = 2\delta / \sqrt{2\ln 2}$, 回波强度 $I = \sqrt{2\pi}A\delta$, 地物到发射器之间的距离 $R = c \cdot (\mu_r - \mu_e)/2$, 目标地物的散射截面 $\sigma = C_{cal}R^4A\delta$ 。其中 δ 为脉冲的标准差, c 为光速, μ_r 为回波脉冲的位置, μ_e 为发射脉冲的位置, A 为脉冲的振幅, C_{cal} 为校正系数。

3 结 论

本文针对弱波与叠加波的检测问题提出了一种改进的 RGD 算法,在检测之前采用高斯滤波降低了

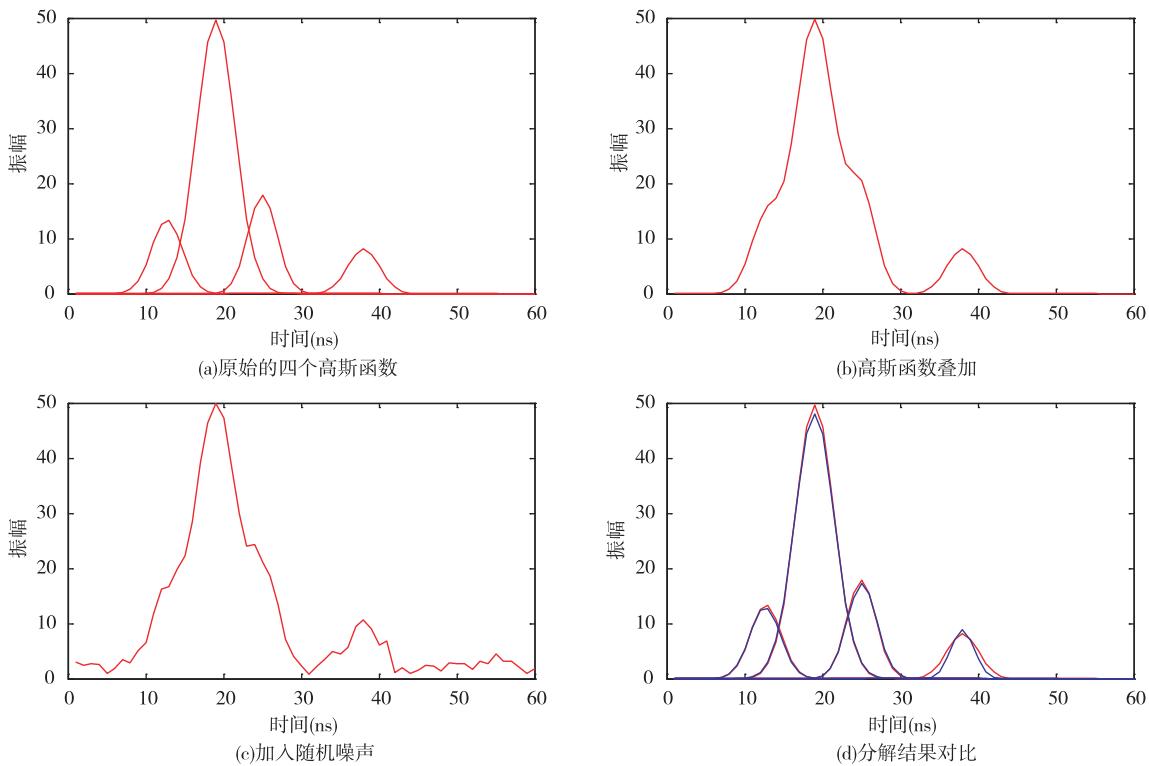


图 8 模拟波形的验证

噪声的影响,之后完成了对全波形数据的分解,最终得到地物的回波波形信息。结合实测数据分解结果的统计分析以及仿真实验,验证了改进算法在叠加波和弱波提取上的有效性和鲁棒性。实验证实,该算法能够有效地检测出复杂波形中的叠加波和振幅值接近阈值水平的弱波,弥补了 COG 算法与 GPF 算法的不足。由于在参数优化时使用了 Trust Region 方法,并采用更为有效的策略来计算增长步,因此分解的结果更加稳健。相比 RGD 算法,本文算法的优势在于提高了信噪比,能够适用于复杂的叠加波形,而且检测不受单次回波叠加个数限制,因此提高了叠加波的识别率。而如何对算法进行完善,使其能够推广应用于商业的生成激光点云的工作中还需作进一步的研究。

参考文献

- [1] Zhu J F,Zhang Z X,Hu X Y et al. Analysis and application of lidar waveform data using a progressive waveform decomposition method. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*,2001,XXXVIII-5/W12:31-36
- [2] Ma H C,Li Q. Modified EM algorithm and its application to the decomposition of laser scanning waveform data. *Journal of Remote Sensing*,2009,13(1):35-41
- [3] Hofton M A,Minster J B,Blair J B. Decomposition of laser altimeter waveforms. *IEEE Transactions on Geoscience and Remote Sensing*,38(4):1989-1996
- [4] Chauve A,Mallet C,Bretar F,et al. Processing full-waveform lidar data: modelling raw signals. *International Archives of Photogrammetry,Remote Sensing and Spatial Information Sciences*,2007,XXXVI(Part 3/W52):102-107
- [5] Roncat A,Wagner W,Melzer T,et al. Echo detection and localization in full-waveform airborne laser scanner data using the averaged square difference function estimator. *The Photogrammetric Journal of Finland*,2008,21(1):62-75
- [6] Roncat A,Bergauer G,Pfeifer N. B-spline deconvolution for differential target cross-section determination in full-waveform laser scanning data. *Journal of Photogrammetry and Remote Sensing*,2011,(66):418-428
- [7] Stilla U,Yao W,Jutzi B. Detection of weak laser pulses by full waveform stacking. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*,2007,36(3/W49A):25-30
- [8] Yu C L,Mills J P,Voysey S S. Rigorous pulse detection from fullwaveform airborne laser scanning data. *International Journal of Remote Sensing*,2010,31(5):1303-1324
- [9] Wagner W,Ullrich A,Ducic V. Gaussian decomposition

- and calibration of a novel small-footprint full-wave form digitising airborne laser scanner. *Journal of Photogrammetry & Remote Sensing*, 2006, (60):100-112
- [10] Jie S, Charles K T. Topographic Laser Ranging and Scanning Principles and Processing. Boca Raton: CRCPress, 2008. 223-227
- [11] Wagner W, Ullrich A, Melze T, et al. From single-pulse to full-waveform airborne laser scanners: potential and practical challenges. *International Archives of Photogrammetry and Remote Sensing*, 2004, XXXV (Part B3): 201-206
- [12] Jutzi B, Still U. Laser pulse analysis for reconstruction and classification of urban objects. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2003, XXXIV (Part 3/W8): 151-156
- [13] 王燕. 二阶导数的五点数值微分公式及外推算法. 天津理工大学学报, 2009, 25(4): 37-39
- [14] Persson Å, Söderman U, Töpel J, et al. Visualization and analysis of full-waveform airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2005, 36 (Part 3/W19): 103-108
- [15] Nocedal J, Wright S J. Numerical Optimization. London: Springer, 1999. 65-84
- [16] More J J, Sorensen D C. Computing a trust region step. *SIAM J Sci Stat COMPUT*, 1983, 4(3): 553-572

Research on decomposition of full-waveform airborne laser data

Liu Zhao*, Zhang Aiwu*, Duan Yihao*, Wang Shumin**

(* Ministry of Education Key Laboratory of 3D Information Acquisition & Application,
Capital Normal University, Beijing 100048)

(** Institute of Earthquake Science, China Earthquake Administration, Beijing 100036)

Abstract

Based on the analysis of the present techniques of decomposing full-waveform airborne laser data, a modified rigorous Gaussian detection (RGD) algorithm was proposed to solve the problem that in waveform decomposition the center of gravity (COG) method and the Gaussian pulse fitting (GPF) method can not extract weak pulses and overlapped pulses in the full-waveform data, so some earth surface feature information can not be achieved. Firstly, the algorithm uses the Gaussian filtering to improve the signal to noise ratio. Then initial parameters are estimated through pulse detection, and the trust region parameters optimization method is adopted to perform the Gaussian fitting. Finally, the weak pulses and the overlapped pulses are effectively decomposed. By using the model parameters such as standard deviation, amplitude and the position of each single echo extracted from decomposition, the full width at half maximum (FWHM), backscattered cross-section and echo intensity, etc., can be obtained. The numerous experiments were performed, and the results show that the proposed algorithm can effectively improve the recognition rate of overlapped pulses and reduce the error of waveform fitting compared with the rigorous Gaussian detection algorithm.

Key words: full-waveform, weak pulses and overlapping pulses, decomposition of waveform, Gaussian detection (RGD) algorithm, Trust Region