

文章编号 1004-924X(2018)01-0161-11

## 偏正态全波激光雷达数据的可变分量波形分解

赵泉华\*, 陈为多, 王 玉, 李 玉

(辽宁工程技术大学 测绘与地理科学学院, 辽宁 阜新 123000)

**摘要:**针对传统方法难以实现全波激光雷达数据中非对称波形分解的问题,本文提出一种结合可变分量偏正态模型和可逆跳马尔科夫链蒙特卡洛(RJMCMC)算法的波形分解方法。首先,利用能量函数刻画服从偏正态分布的理想波形与实际波形间的差异程度,并用 Gibbs 分布定义其似然函数;然后,定义理想波形参数模型的先验分布;在贝叶斯定理框架下,建立具有分量可变性的波形分解模型;设计 RJMCMC 的移动操作,确定偏正态分布中的分量数以及求解模型参数。利用提出算法,对不同波形特征(偏态、正态)的 ICESat-GLAS 全波激光雷达数据进行可变分量分解实验。实验结果表明:实验波形结果与实际波形数据相关系数达到 0.989 以上,所提方法不仅能够同时实现对偏态数据和正态数据的拟合,还能更为准确地确定波形分量数。证明了该方法能实现全波激光雷达数据的精确分解,且分解结果与对应地物高程信息相符。

**关键词:**全波激光雷达;波形分解;偏正态分布;RJMCMC 算法

**中图分类号:**TP751 **文献标识码:**A **doi:**10.3788/OPE.20182601.0161

## Full-waveform LiDAR data decomposition based on skew-normal distribution with unknown number of components

ZHAO Quan-hua\*, CHEN Wei-duo, WANG Yu, LI Yu

(School of Geomatics, Liaoning Technical University, Fuxin 123000, China)

\* Corresponding author, E-mail: zhaquanhua@lntu.edu.cn

**Abstract:** To decompose asymmetric full-waveform LiDAR data with unknown number of components, a full-waveform LiDAR decomposition method was proposed based on skew-normal distribution and reversible-jump Markov Chain Monte Carlo (RJMCMC) algorithm, which can automatically determine the numbers of components. First, the energy function was used to describe the differences between the actual waveform and the ideal waveform that obeyed the skew-normal distribution, and the likelihood function was defined by Gibbs distribution. Second, the parameter models of the ideal waveform were established using the prior distribution. Then the Bayesian paradigm was followed to build the ideal waveform model. Third, an RJMCMC algorithm was designed to determine the numbers of components and decompose the waveform. The proposed algorithm was used to decompose ICESat-GLAS waveform data in various typical regions. Experimental results indicate that the cross-correlation of the true data and the result is up to 98.9%.

收稿日期:2017-05-15;修订日期:2017-07-20.

基金项目:辽宁省教育厅科学技术研究一般资助项目(No. LNCL009);国家自然科学基金青年基金资助项目(No. 41301479)

The proposed method can not only fit the skewed waveform data and normal waveform data, but also more accurately determine the number of components in comparison to other methods. It can realize the accurate decomposition of full-waveform LiDAR data, and the decomposition result is consistent with the corresponding elevation information.

**Key words:** full-waveform LiDAR; waveform decomposition; skew-normal distribution; RJMCMC algorithm

## 1 引言

相较于传统激光雷达系统仅能够记录有限个离散的返回信号<sup>[1]</sup>,全波激光雷达(Full-waveform LiDAR)能够记录更为完整的波形数据,而波形分解是全波 LiDAR 数据处理的基础。通过波形分解可还原单回波信息,为生成高密度点云数据提供了基础。对全波 LiDAR 进行建模可以得到更多信息参数,如位置、振幅和脉冲宽度信息能够反映地物高程信息和背向散射特性,这些信息可以有效辅助地物辨识与分割。波形分解是用理想波形拟合实际波形的过程,其中包括两个任务:构建理想模型以及求解模型参数。

目前,已经有许多学者对全波 LiDAR 数据的理想模型进行了研究。Wagner<sup>[2]</sup>等针对 RIEGL LMS-Q560 机载 LiDAR 脉冲波形的特征,指出由雷达方程推导出多个高斯函数可拟合全波 LiDAR 波形数据。以此为依据,高斯模型被广泛用于解决全波 LiDAR 数据的波形分解问题<sup>[3-4]</sup>。作为经典模型,高斯模型虽适用于大部分全波 LiDAR 数据的波形分解,但并不适于波形呈偏态分布的波形分解。Chauve<sup>[5]</sup>等在改进高斯模型的基础上,提出了基于对数高斯分布和广义高斯分布的分解模型。对数高斯分布仅适用于拟合具有偏态性的波形数据,但无法有效拟合对称波形;虽然广义高斯分布可以提高对称波形的拟合精度,却不能拟合复杂的偏态波形。因此,上述统计分布作为理想模型均具有局限性。

在波形分解中,常用的模型参数求解方法包括:基于非线性最小二乘原理的 LM (Levenberg Marquardt) 算法<sup>[6-7]</sup>、高斯-牛顿法和基于极大似然原理的 EM (Expectation Maximum) 算法<sup>[8-9]</sup>等,但这些方法需要人为确定波形分量数和模型初始参数。错误的波形分量和初始参数设置将导致波形分解结果陷入局部最优解,且当波形分量发生混叠时很难人为地确定波形分量数。

全波 LiDAR 系统采样整个光斑内反射能

量,以获取同一高程地物的反射脉冲强度。大光斑 LiDAR 系统的光斑直径一般为 8~70 m<sup>[10]</sup>,相较于小光斑其光斑覆盖面积较大。光斑内地物类别和地形特征的多样性均可导致该区域高程分布复杂。当光斑内包含多类地物时,全波 LiDAR 数据中记录的同一高程反射脉冲强度可能来自具有不同反射特性的各类地物目标,从而导致波形数据呈偏态。当光斑内仅有单一地物时,以大片森林为例,大光斑 LiDAR 发射的激光脉冲能穿透森林冠层,波形数据能够反映冠层在高程方向的分布情况,混杂的树冠反射特性也会导致波形的偏态性<sup>[11]</sup>。庞勇<sup>[12]</sup>等证明了地形坡度对大光斑 LiDAR 波形影响较大,即使微弱的地形起伏也会造成明显的地面高程差,当地形起伏大于一定坡度时,下坡的高地物回波与上坡的地面回波会发生混叠。

综上所述,传统的具有固定分量数的对称性高斯混合模型并不普遍适用于大光斑全波 LiDAR 数据的波形分解。因此,本文提出可变量数数的偏正态分布混合模型,以实现大光斑全波 LiDAR 数据的波形分解。首先,建立全波 LiDAR 数据理想波形的偏正态分布混合模型,并利用能量函数刻画理想波形与实际波形的差异程度,用 Gibbs 分布定义其似然函数;然后,根据贝叶斯定理<sup>[13]</sup>,建立基于可变量的波形分解模型;再利用 RJMCMC 算法模拟该模型。为了验证提出算法的可行性及有效性,以 ICESat-GLAS 大光斑全波 LIDAR 数据为实验数据,对其进行可变量数波形分解实验。通过实验结果看出,提出方法不仅能够自动确定波形分解数,还能同时实现对高斯模型和偏态模型的拟合,且分解得到的波形信息与实际地物相对应。

## 2 算法描述

全波 LiDAR 数据为各采样点反射能量的集合,可表示为  $\mathbf{z} = \{(\mathbf{x}, \mathbf{y})\}$ 。其中,  $\mathbf{x} = \{x_k; k = 1, 2, \dots, m\}$  为在  $[0, N]$  区间内采样点的横坐标集

合,代表采样的相对时间; $\mathbf{y}=\{y_k; k=1,2,\dots,m\}$ 为采样点的反射能量集合; $m$ 为有效采样点总数。

### 2.1 可变量偏正态波形分解模型

在波形分解中,需定义全波 LiDAR 数据的理想波形并以此拟合实际波形。设理想波形可由  $n$  个偏正态分布加权和表示,则其概率密度函数可表示为:

$$\tilde{y}_k = \sum_{i=1}^n \left[ \omega_i \frac{2}{\sigma_i} \varphi \left( \frac{x_k - \mu_i}{\sigma_i} \right) \Phi \left[ \alpha_i \frac{x_k - \mu_i}{\sigma_i} \right] \right], \quad (1)$$

式中: $\omega_i$ 为第  $i$  个偏正态分布的权重,各分量权重和为 1; $\mu_i, \sigma_i$  及  $\alpha_i$  分别代表第  $i$  个偏正态分布的均值、方差及偏度值,当偏度值为 0 时偏正态分布为高斯分布; $\varphi(t)$  和  $\Phi(t)$  分别代表标准正态分布的概率密度函数和分布函数,其函数表达式分别为:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{t^2}{2} \right\}, \quad (2)$$

$$\Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{t^2}{2} \right\} dt. \quad (3)$$

定义  $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}, n\}$  为所有参数构成的矢量集合,其中  $\boldsymbol{\mu} = \{\mu_i, i=1, \dots, n\}$ ,  $\boldsymbol{\sigma} = \{\sigma_i, i=1, \dots, n\}$ ,  $\boldsymbol{\alpha} = \{\alpha_i, i=1, \dots, n\}$ ,  $\boldsymbol{\omega} = \{\omega_i, i=1, \dots, n\}$ 。本文利用能量函数  $U(\mathbf{z} | \boldsymbol{\theta})$  定义理想波形  $\tilde{y}_k$  和实际波形  $y_k$  的拟合程度,并构建其 Gibbs 分布,可表示为:

$$L(\mathbf{z} | \boldsymbol{\theta}) = \frac{1}{\beta} \exp(-U(\mathbf{z} | \boldsymbol{\theta})) = \frac{1}{\beta} \exp \left( - \sum_{k=1}^m |y_k - \tilde{y}_k| \right), \quad (4)$$

式中  $\beta$  为归一化常数。

假设  $\mu_i (i=1, 2, \dots, n)$  服从区间  $[1, N]$  均匀分布且相互独立;标准差  $\sigma_i (i=1, 2, \dots, n)$  和偏度值  $\alpha_i (i=1, 2, \dots, n)$  分别服从均值为  $\mu_\sigma, \mu_\alpha$  和标准差为  $\sigma_\sigma, \sigma_\alpha$  的正态分布且相互独立;权重  $\omega_i (i=1, 2, \dots, n)$  服从  $[0, 1]$  上均匀分布且相互独立;分量的个数  $n$  服从均值为  $\lambda$  的泊松分布,则各参数的先验概率模型可表达为:

$$p(\boldsymbol{\mu}) = p(\mu_1, \mu_2, \dots, \mu_n) = \prod_{i=1}^n p(\mu_i) = \frac{1}{N^n}, \quad (5)$$

$$p(\boldsymbol{\sigma}) = p(\sigma_1, \sigma_2, \dots, \sigma_n) = \prod_{i=1}^n p(\sigma_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_\sigma} \exp \left\{ -\frac{(\sigma_i - \mu_\sigma)^2}{2\sigma_\sigma^2} \right\}, \quad (6)$$

$$p(\boldsymbol{\alpha}) = p(\alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{i=1}^n p(\alpha_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_\alpha} \exp \left\{ -\frac{(\alpha_i - \mu_\alpha)^2}{2\sigma_\alpha^2} \right\}, \quad (7)$$

$$p(\boldsymbol{\omega} | m) = p(\omega_1, \omega_2, \dots, \omega_n) = \prod_{i=1}^n p(\omega_i), \quad (8)$$

$$p(n) = \frac{\lambda^n}{n!} \exp(-\lambda), \quad (9)$$

式中: $0 < \mu_1 < \mu_2 < \dots < \mu_n < N$ ,  $\mu_\sigma, \sigma_\sigma, \mu_\alpha, \sigma_\alpha$  和  $\lambda$  为预先自定义参数。

根据贝叶斯定理可得在已知全波 LiDAR 数据  $\mathbf{z}$  条件下的参数  $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}$  及  $n$  的联合后验概率密度函数:

$$p(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}, n) \propto L(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}, n) p(\boldsymbol{\mu} | n) p(\boldsymbol{\sigma} | n) p(\boldsymbol{\alpha} | n) p(\boldsymbol{\omega} | n) p(n), \quad (10)$$

即为可变量偏正态波形的分解模型。

### 2.2 可变量偏正态波形分解模型求解

可变量偏正态波形分解模型建立完成后,利用 RJMCMC 算法模拟该模型,以自动确定波形数据的最优分解数。假设参数矢量中的各参数相互独立,按式(10)采用 RJMCMC 算法进行采样。在每次采样过程中,当参数维度未发生变化时,选参数矢量  $\boldsymbol{\theta}^*$  的接受率可表示为:

$$a(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = \min \left\{ 1, \frac{p(\mathbf{z} | \boldsymbol{\theta}^*) p(\boldsymbol{\theta}^*)}{p(\mathbf{z} | \boldsymbol{\theta}) p(\boldsymbol{\theta})} \right\}, \quad (11)$$

当参数维度发生变化时,假设候选参数矢量  $\boldsymbol{\theta}^*$  的维度高于  $\boldsymbol{\theta}$ ,其接受率可表示为:

$$a(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = \min \left\{ 1, \frac{p(\boldsymbol{\theta}^* | \mathbf{z}) r(\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta} | \mathbf{z}) r(\boldsymbol{\theta}) q(\mathbf{u})} \left| \frac{\partial(\boldsymbol{\theta}^*)}{\partial(\boldsymbol{\theta}, \mathbf{u})} \right| \right\}, \quad (12)$$

式中: $\mathbf{u}$  为一个随机矢量,可确保状态由  $\boldsymbol{\theta}$  到  $\boldsymbol{\theta}^*$  变化时,维度平衡,即  $|\boldsymbol{\theta}| + |\mathbf{u}| = |\boldsymbol{\theta}^*|$ ;  $q(\mathbf{u})$  是  $\mathbf{u}$  的概率密度函数; $r(\boldsymbol{\theta})$  和  $r(\boldsymbol{\theta}^*)$  分别为状态  $\boldsymbol{\theta}$  和  $\boldsymbol{\theta}^*$  对应的操作概率; $|\partial(\boldsymbol{\theta}^*)/\partial(\boldsymbol{\theta}, \mathbf{u})|$  表示为由状态  $\boldsymbol{\theta}$  到  $\boldsymbol{\theta}^*$  的 Jacobian 项。

在 RJMCMC 算法中,设计了 4 个移动操作,分别为更新偏正态分布参数矢量  $\boldsymbol{\varphi} = \{\varphi_i, i=1, \dots, n\} = \{(\mu_i, \sigma_i, \alpha_i), i=1, \dots, n\}$ 、更新权重  $\boldsymbol{\omega}$ 、生成/删除偏正态分量及合并/分解偏正态

分量。

具体操作如下：

(1)更新偏正态分布参数矢量  $\boldsymbol{\varphi}$ ；

在  $\boldsymbol{\varphi} = \{\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_n\}$  中随机抽取第  $i$  个偏正态分量  $\boldsymbol{\varphi}_i$ ，顺序改变  $\boldsymbol{\varphi}_i$  中的  $\mu_i, \sigma_i$  和  $\alpha_i$ 。以更新均值  $\boldsymbol{\mu}$  为例，其候选均值  $\mu_i^*$  是从均值为  $\mu_i$ 、标准差为  $\epsilon_\mu$  的正态分布中随机抽取得到，则  $\boldsymbol{\mu}^* = \{\mu_1, \dots, \mu_{i-1}, \mu_i^*, \mu_{i+1}, \dots, \mu_n\}$ ；其中  $\epsilon_\mu$  为实验设计均值移动步长，为预先定义参数。更新均值  $\boldsymbol{\mu}$  的接受率为：

$$a(\boldsymbol{\mu}, \boldsymbol{\mu}^*) = \min\{1, R_\mu\} = \min\left\{1, \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(\boldsymbol{\mu}^*)}{p(\boldsymbol{\mu})}\right\}, \quad (13)$$

式中  $\boldsymbol{\theta}^* = \{\boldsymbol{\mu}^*, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}, n\}$ 。

同理分别以移动步长  $\epsilon_\sigma$  和  $\epsilon_\alpha$  更新方差  $\boldsymbol{\sigma}$  和更新偏度值  $\boldsymbol{\alpha}$  的接受率分别为：

$$a(\boldsymbol{\sigma}, \boldsymbol{\sigma}^*) = \min\{1, R_\sigma\} = \min\left\{1, \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(\boldsymbol{\sigma}^*)}{p(\boldsymbol{\sigma})}\right\}, \quad (14)$$

$$a(\boldsymbol{\alpha}, \boldsymbol{\alpha}^*) = \min\{1, R_\alpha\} = \min\left\{1, \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(\boldsymbol{\alpha}^*)}{p(\boldsymbol{\alpha})}\right\}, \quad (15)$$

式(13)和式(14)中  $\boldsymbol{\theta}^*$  分别为  $\{\boldsymbol{\mu}, \boldsymbol{\sigma}^*, \boldsymbol{\alpha}, \boldsymbol{\omega}, n\}$  和  $\{\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}^*, \boldsymbol{\omega}, n\}$ 。

(2)更新权重  $\boldsymbol{\omega}$ ；

首先在  $\boldsymbol{\omega}$  中随机抽取第  $i$  个偏正态分量，然后随机选取第  $i$  个的邻域分量(第  $i-1$  或  $i+1$  个)。以选取第  $i$  个和第  $i+1$  个偏正态分量为例，生成随机数  $l \in (0, 1)$ ，得  $\omega_i^* = (\omega_i + \omega_{i+1}) \times l$  和  $\omega_{i+1}^* = (\omega_i + \omega_{i+1}) \times (1-l)$ ，则  $\boldsymbol{\omega}^* = \{\omega_1, \dots, \omega_i^*, \omega_{i+1}^*, \dots, \omega_n\}$ 。更新权重  $\boldsymbol{\omega}$  为  $\boldsymbol{\omega}^*$  的接受率为：

$$a(\boldsymbol{\omega}, \boldsymbol{\omega}^*) = \min\{1, R_\omega\} = \min\left\{1, \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(\boldsymbol{\omega}^*)}{p(\boldsymbol{\omega})}\right\}, \quad (16)$$

式中  $\boldsymbol{\theta}^* = \{\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\omega}^*, n\}$ 。

(3)生成/删除偏正态分量；

生成与删除偏正态分量操作是对偶操作，下面以生成操作为例进行说明。首先随机选取第  $i$  个偏正态分量，然后以等概率选择向左或向右生成偏正态分量。以选择向右生成为例，则  $\mu_i^*$  服

从在  $(\mu_i, \mu_{i+1})$  的均匀分布， $\sigma_i^*$ 、 $\alpha_i^*$ 、 $\omega_i^*$  按生成候选参数方法生成，得到新的参数集合  $\boldsymbol{\theta}^* = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_i, \boldsymbol{\theta}_i^*, \boldsymbol{\theta}_{i+1}, \dots, \boldsymbol{\theta}_n\}$ 。则生成操作的接受率为：

$$a_b = \min\{1, R_b\}, \quad (17)$$

$$R_b = \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(m+1)}{p(m)} \times \frac{r_b}{r_d}, \quad (18)$$

式中： $r_b = b_i/n$  和  $r_d = d_i/n+1$  分别为生成高斯分量和删除高斯分量的总概率， $b_i$  和  $d_i$  分别为随机选择的第  $i$  个偏正态分量进行生成和删除操作的概率。当  $i=1$  时， $d_1=0, b_1=1$ ；当  $i=n$  时， $b_n=0, d_n=1$ ；当  $i=2, 3, \dots, n-1$  时， $b_i = d_i = 0.5$ 。

由于删除与生成互为对偶操作，则删除偏正态分量的接受率为：

$$a_d = \min\{1, 1/R_b\}. \quad (19)$$

(4)合并/分裂偏正态分量。

合并与分裂操作是对偶操作，下面以分裂操作为例进行说明。首先随机选取第  $i \in \{1, 2, \dots, n\}$  个偏正态分量，然后以等概率选择向左或向右分裂偏正态分量。以选择向右分裂为例，则  $\mu_i^*$  和  $\mu_{i+1}^*$  服从在  $(\mu_{i-1}, \mu_{i+1})$  的均匀分布， $\sigma_i^*$ 、 $\sigma_{i+1}^*$ 、 $\alpha_i^*$ 、 $\alpha_{i+1}^*$ 、 $\omega_i^*$ 、 $\omega_{i+1}^*$  按生成候选参数方法生成，得到新的参数集合  $\boldsymbol{\theta}^* = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i^* \setminus, \boldsymbol{\theta}_{i+1}^*, \dots, \boldsymbol{\theta}_n\}$ 。则分裂操作的接受率为：

$$a_s = \min\{1, R_s\}, \quad (20)$$

$$R_s = \frac{L(\mathbf{z}|\boldsymbol{\theta}^*)}{L(\mathbf{z}|\boldsymbol{\theta})} \times \frac{p(m+1)}{p(m)} \times \frac{r_s}{r_m}, \quad (21)$$

式中： $r_s = s_i/n$  和  $r_m = m_i/n+1$  分别为生成偏正态分量和删除偏正态分量的总概率， $s_i$  和  $m_i$  分别为随机选择的第  $i$  个偏正态分量进行分裂和合并操作的概率。当  $i=1$  时， $m_1=0, s_1=1$ ；当  $i=n$  时， $s_n=0, m_n=1$ ；当  $i=2, 3, \dots, n-1$  时， $s_i = m_i = 0.5$ 。

合并与分裂互为对偶操作，则合并偏正态分量的接受率为：

$$a_m = \min\{1, 1/R_s\}. \quad (22)$$

### 3 实验与分析

提出方法假设全波 LiDAR 数据的理想波形服从偏正态分布，偏正态分布混合模型能同时实现偏态模型和高斯模型的拟合，故实验分别设计

对偏态波形数据和正态波形数据进行可变量波形分解。

实验得到的偏正态分量均值为对应地物的采样相对时间,根据  $d(\text{距离}) = c(\text{光速}) \times t(\text{相对时间})$  可得各地物的相对距离。实验中取地面高程为 0,则通过偏正态分量均值可推算各地物对地高程。

实验数据采用 ICESat-GLAS 大光斑全波 LiDAR 数据,分别拟合并分析偏态波形数据和正态波形数据。图 1、图 2 分别为偏态波形数据和正态波形数据对应的实验区域的光学遥感图像。通过分解得到的分量数、拟合精度和推算出的各地物相对高程 3 个方面,验证提出方法的有效性和准确性。

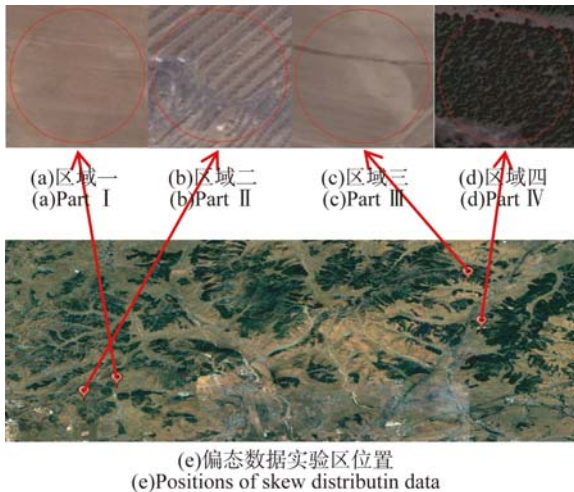


图 1 偏态数据实验区  
Fig. 1 Study areas of skew distribution data

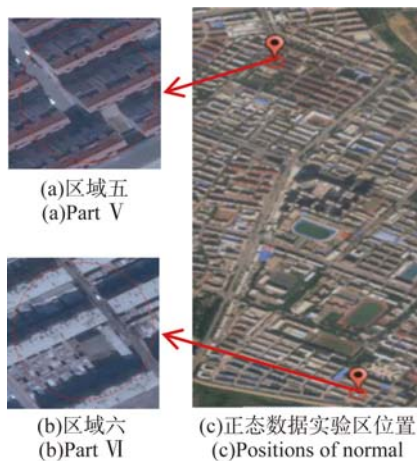


图 2 正态数据实验区  
Fig. 2 Study areas of gaussian distribution data

#### 4 实验结果与讨论

针对地形起伏大或垂直结构复杂的区域进行研究,以验证提出方法对偏态分布数据分解的可行性及优势。如图 1 所示,区域一到区域四分别位于图 1(e)中的 4 个光标位置,经实地考察验证光斑均位于辽宁省阜新市阜蒙县,数据采集时间分别为 2006 年 5 月、2006 年 2 月、2003 年 9 月和 2003 年 9 月。图 1(a)~(d)对应区域分别为山坡裸地、山坡较高农作物区域、山坡低矮农作物区域和高大树林区域,其中,图 1(a)~(b)分别为相同山坡的裸地和耕地,光斑地表粗糙且高程差达到 7 m;区域二每年 2 月会残留大量玉米杆,玉米杆成整齐的带状分布,其高度约为 1.2 m;图 1(c) 为大片密集花生地,光斑内地表高程差达到 8.2 m;图 1(d)为平坦地区的茂密杨树林。

4 个实验区域对应的实际波形数据及其对应的分解结果见图 3(彩图见期刊电子版),其中原始数据的横坐标为采样时间,实验结果的横坐标为由采样相对时间转换成的采样相对高程,纵坐标为反射的脉冲强度。图 3(a1)~(a3)为区域一的实验结果,仅有一个地面反射的波形信号返回且回波形状左偏,这与区域一的山坡粗糙裸地的地貌相符;图 3(b1)~(b3)结果显示区域二返回两次回波:根据接收时间的先后顺序分别确定为玉米杆回波和地面回波,其中地面回波的偏度较大,由偏正态分量的均值参数推算出回波对应地物的高程相差 1.05 m;图 3(c1)~(c3)为区域三山坡花生地回波信号分解结果。根据接收时间的先后顺序分别确定为花生苗回波和地面回波,花生苗高度较小且覆盖地表密集导致花生苗回波脉冲强度较高,且随地表坡度变化呈现偏态,结果显示花生植被约高 0.15 m;图 3(d)所在区域四在 6 月期间生长着大面积茂盛的杨树,由于大光斑激光雷达能部分地穿透植被冠层,波形数据能够反映光斑内树木垂直方向分布情况,茂密树林区域的回波形状呈现偏态,且树木枝叶茂密区(反射能量最多)距地高度约为 4.95 m。

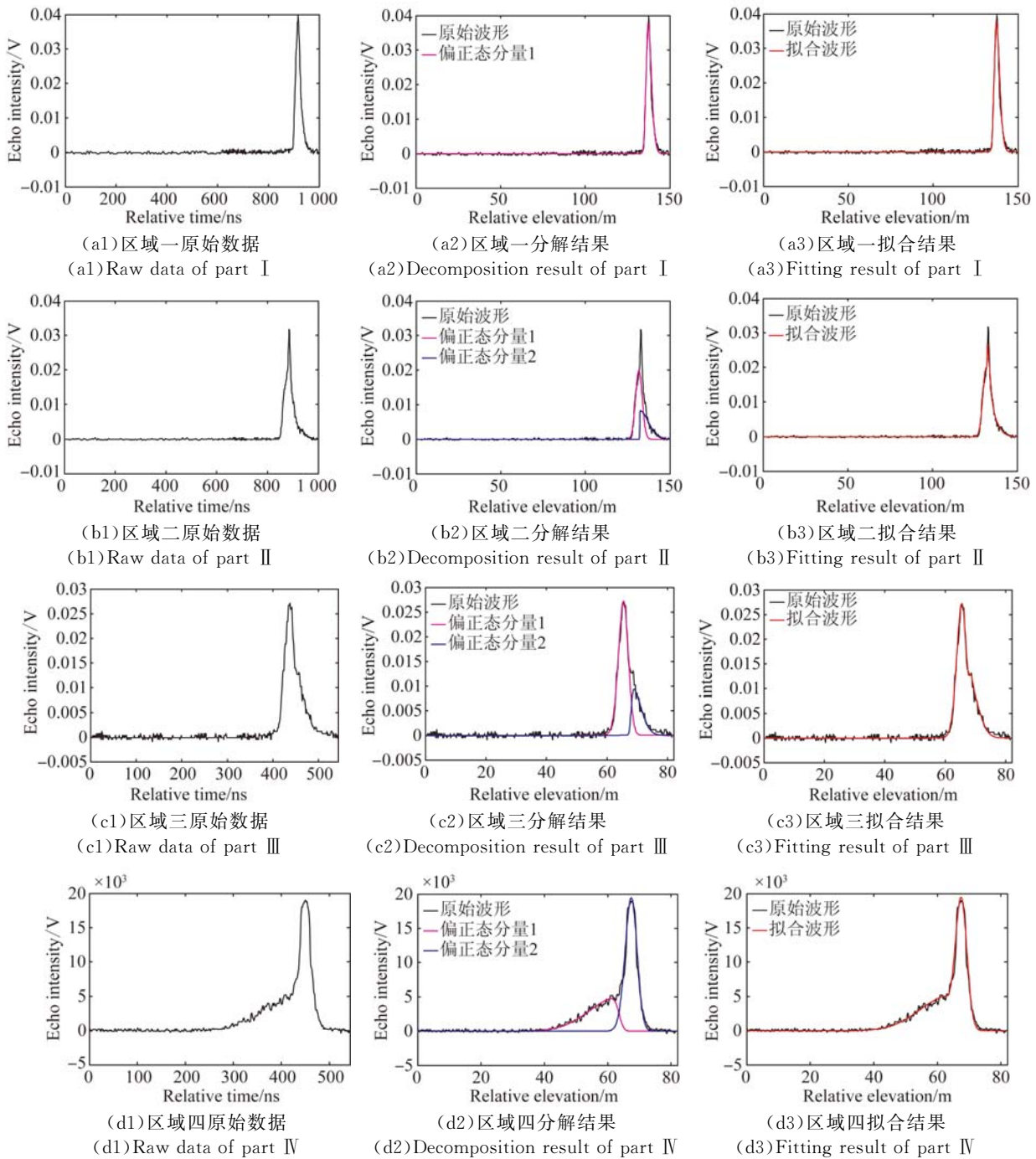


图 3 偏态数据分解结果

Fig. 3 Decomposition results of skew distribution data

为了验证提出方法能够实现对偏态波形数据分解的同时,对正态分布数据分解同样具有可行性,本文选取两处地势平坦且垂直结构单一地区的波形数据进行分解实验。图 2 中两处波形数据采集时间为 2004 年 10 月,地点为辽宁省阜新市市区。经实地勘测区域五为带有楼顶天台的居民楼,天台与楼顶相对地面高程为 15.7 m 和 20.1 m,而

区域六为复杂楼群,其中包括一层煤棚,二层商用建筑,四层居民楼和五层居民楼,实地勘测的相对高程为:2.5 m,7.4 m,12.0 m 和 15.1 m。图 4 (彩图见期刊电子版)为实验区域对应的实际波形及其分解结果,其中,图 4(a1)~(a2)为区域五的波形分解结果,分解出 3 个分量,分别代表地面、天台和楼顶,其相对高程分别为 0 m,16.05 m 和

19.5 m;图 4(b1)~(b2)为区域六的波形分解结果,分解出 5 个分量,其相对高程为 0 m,2.25 m,6.75 m,11.4 m 和 14.85 m,该区域结果局部放大图显示本文算法可精确拟合细微的回波信息。

通过图 4 可以看出,提出方法不仅可以准确地确定正态分布数据的波形数,还可以较好地实现波形数据的拟合,从而验证提出方法对正态分布数据分解的可行性。

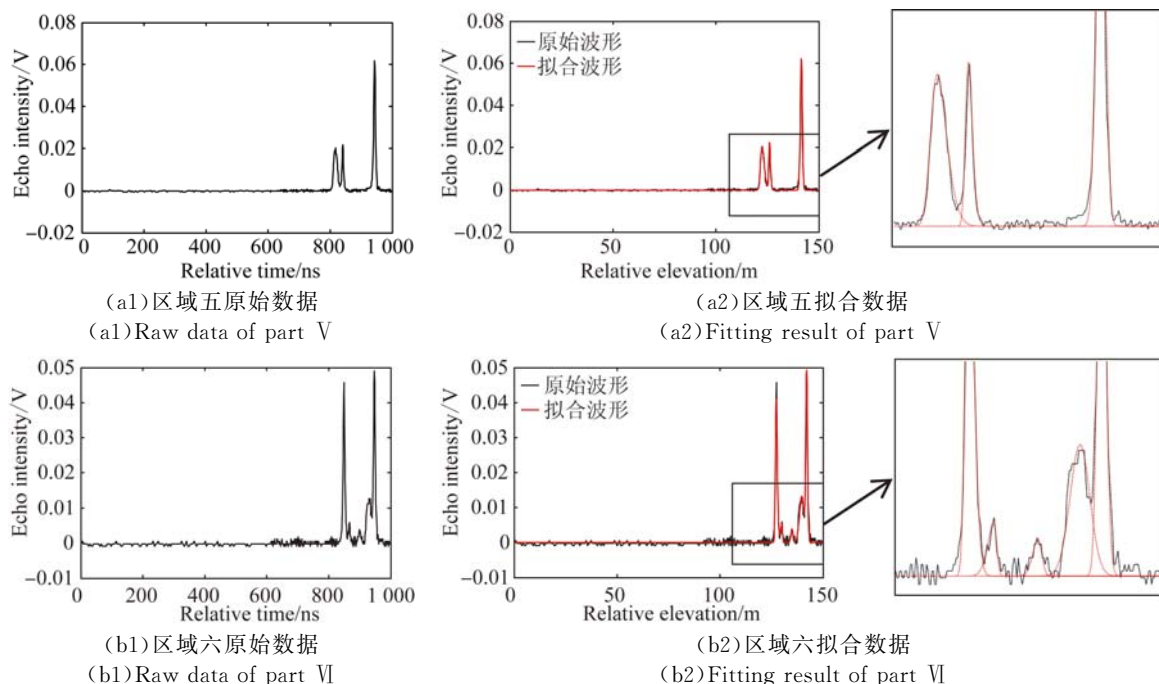


图 4 正态数据分解结果

Fig. 4 Decomposition results of Gaussian distribution data

为验证本文方法拟合全波 LiDAR 数据的精确性,利用相关系数来衡量实验结果与实际数据的近似程度,相关系数越接近 1,说明拟合的效果越好。相关系数公式如下:

$$\rho = \frac{\sum_{k=1}^m (y_k - \bar{y})(\tilde{y}_k - \bar{\tilde{y}})}{\sqrt{\sum_{k=1}^m (y_k - \bar{y})^2} \sqrt{\sum_{k=1}^m (\tilde{y}_k - \bar{\tilde{y}})^2}}, \quad (23)$$

其中  $\bar{y}$  和  $\bar{\tilde{y}}$  分别为实际数据和理想模型的能量均值。计算结果如表 1 所示,相关系数均在 0.989 以上,证明了本文方法在波形拟合上的精确性。

表 1 相关系数

Tab. 1 Cross-correlation coefficient

实验区域	区域一	区域二	区域三	区域四	区域五	区域六
相关系数	0.989 7	0.994 6	0.994 3	0.996 0	0.992 4	0.990 3

### 3.2 算法比较

选择 MATLAB 提供的 CF TOOL 高斯拟合

法为对比方法,与本文提出的方法进行比较,从而验证提出方法的优越性。该方法需人为给定分量个数,图 5(彩图见期刊电子版)为对比方法的分解结果,其中,图 5(a1)~(a2)为区域一波形分解结果,其中图 5(a1)为分解为一个波形分量的结果,该分解结果中分量数虽与实际地物数相符,但结果右尾端不能与实际波形精确拟合;图 5(a2)为分解为两个波形分量的结果,该结果虽能精确拟合波形但分量数多于实际地物数;图 5(b1)~(b2)所示的区域二的分解结果同样存在上述问题,即分解结果不能同时满足拟合的精确性和分量数的合理性;图 5(c)为区域三波形分解结果,虽然分量数与拟合精度达到要求,但推算出花生高为 2.46 m,这与实际情况严重不符;图 5(d)对应的区域四为地形平坦的杨树林,其中树冠与地面在高程方向上分割明显,而波形结果显示树冠回波与地面回波产生了大面积重叠,这与实际情况严重不符;图 5(e)所在的区域五地形平坦且楼

群结构简单,提出方法与对比方法分解结果一致;图 5(f)为区域对比算法分解结果,其中部分细微的回波未能精确拟合,图 5(f)的局部放大图与图 4(b2)的局部放大图对比可证明本文算法的优越性。表 2 为两种方法反演的地物高程数据与真实高程数据对比结果,结果表明本文算法得到的高程数据更为准确。根据波形分解结果和地物相对高程反演结果可以看出,对比方法以下存在三类问题,分别为:拟合精度低;地物数错误;地物高程值错误。

此外,特别针对偏态波形数据的分解,本文引

用赵泉华<sup>[14]</sup>等人的基于高斯混合模型的全波 LiDAR 数据分解方法进行对比实验,该方法能够自动确定分量个数,对比实验的最优分解结果如图 6 所示。其中,图 6(a)分解得到两个波形分量,分量数多于实际地物数;图 6(b)~(c)所示的区域二与区域三分解结果的分量数虽与实际地物相符,但反演得到的地物相对高程信息不正确;图 6(d)的结果不仅分解精度较低,且各波形分量分布与实际地物情况不符。表 3 为两种方法反演的地物高程数据与真实高程数据对比结果,及对比实验结果的错误类型标注。

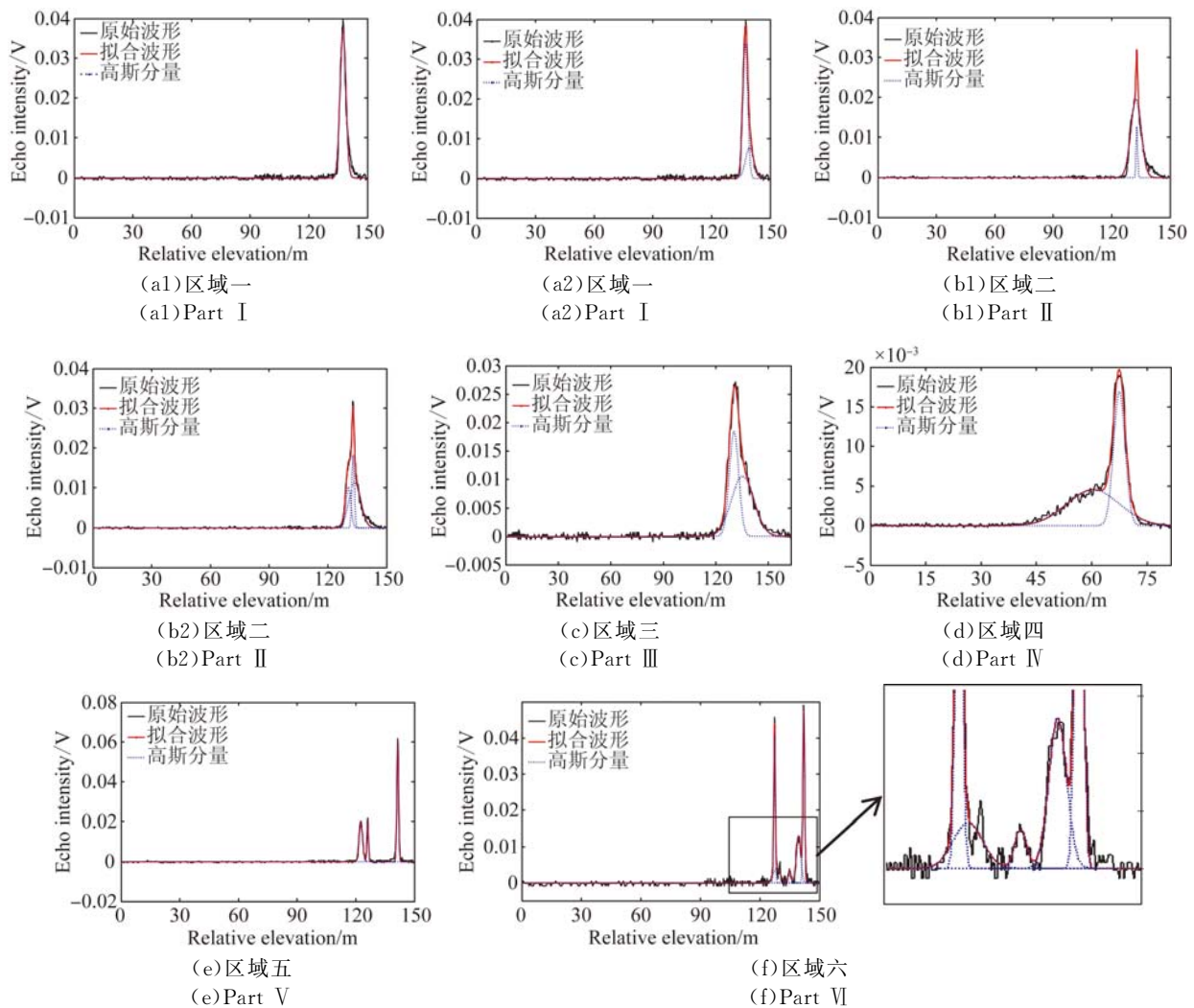


图 5 CFTool 的分解结果

Fig. 5 Decomposition results from CFTool

表 2 2 种方法的对比结果

Tab. 2 Comparison results between two methods

		1	2/m	3/m	4/m	5/m	错误类别
区域一	实际数据	0	—	—	—	—	—
山坡裸地	本文方法	0	—	—	—	—	无
	CFTOOL 方法	0	—	—	—	—	①
区域二	实际数据	0	2.19	—	—	—	②③
	本文方法	0	约 1.2	—	—	—	无
山坡玉米地	本文方法	0	1.05	—	—	—	无
	CFTOOL 方法	0	0.885	3.3	—	—	①③
区域三	实际数据	0	2.19	—	—	—	②③
	本文方法	0	约 0.25	—	—	—	—
山坡花生地	本文方法	0	0.15	—	—	—	无
	CFTOOL 方法	0	2.46	—	—	—	③
区域四	实际数据	0	约 5.4	—	—	—	—
	本文方法	0	4.95	—	—	—	无
平地杨树林	本文方法	0	4.95	—	—	—	无
	CFTOOL 方法	0	7.09	—	—	—	①③
区域五	实际数据	0	15.7	20.1	—	—	—
	本文方法	0	16.05	19.5	—	—	无
平地居民楼	本文方法	0	16.05	19.5	—	—	无
	CFTOOL 方法	0	15.345	18.87	—	—	无
区域六	实际数据	0	2.5	7.4	12.0	15.1	—
	本文方法	0	2.225	6.75	11.4	14.85	无
平地楼群	本文方法	0	2.225	6.75	11.4	14.85	无
	CFTOOL 方法	0	2.595	7.155	13.605	15.715	①③

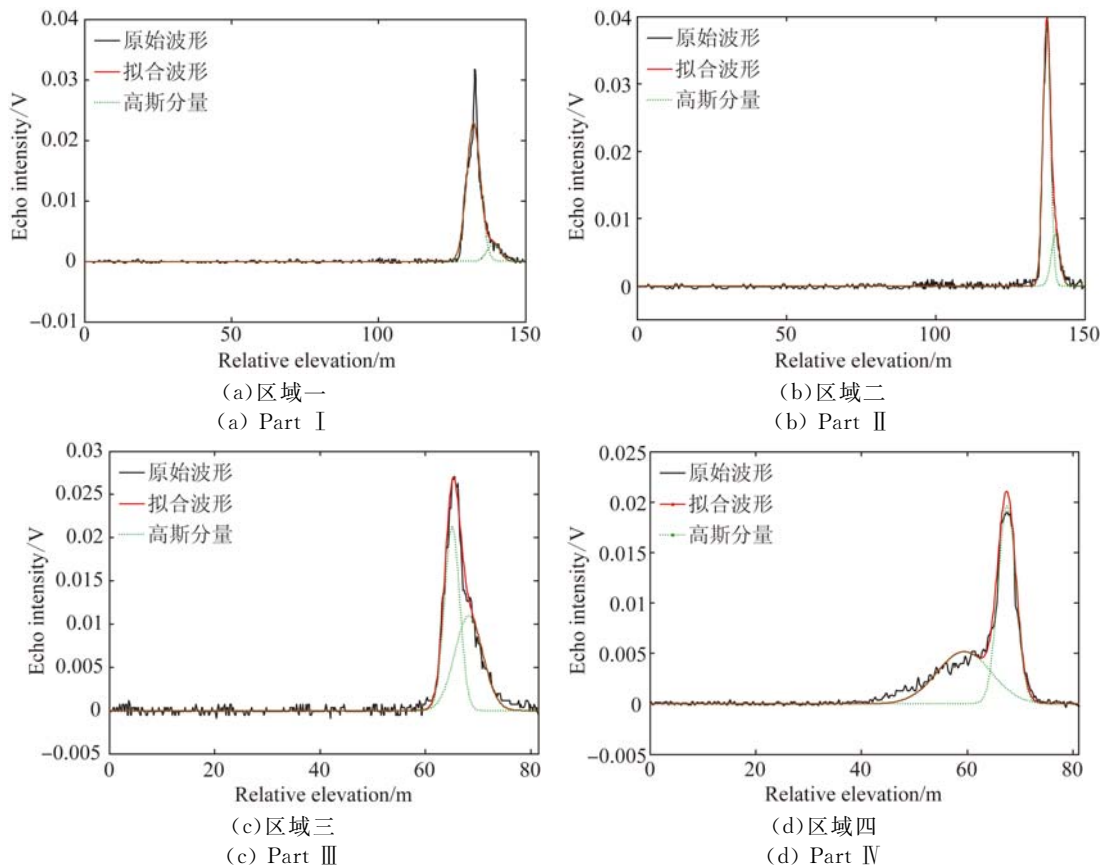


图 6 基于高斯混合模型的分解结果

Fig. 6 Decomposition results based on Gaussian mixture model

表 3 2 种方法的对比结果

Tab. 3 Comparison of results between methods

方法	区域一 山坡裸地			区域二 山坡玉米地			区域三 山坡花生地			区域四 平地杨树林		
	1	2/m	错误类别	1	2/m	错误类别	1	2/m	错误类别	1	2/m	错误类别
实际数据	0	—	—	0	约 1.2	—	0	约 0.25	—	0	约 5.4	—
本文方法	0	—	无	0	1.05	无	0	0.15	无	0	4.95	无
对比方法	0	6.4	②③	0	2.85	③	0	3.15	③	0	9.3	①③

## 4 结 论

针对非对称全波数据分解问题,提出一种结合 RJMCMC 算法和偏正态分布的可变分量波形分解方法。利用提出方法对不同地物和地形区域的 ICESat-GLAS 全波数据进行波形分解实验,实验结果表明提出方法不仅可以自动确定研究区域的波形分解数,还可以较好地实现大坡度地区

和垂直方向结构复杂(如森林)地区波形数据的偏态拟合,同时提高了正态复杂回波的拟合精度,分解结果与实际数据相关系数达到 0.989 以上,证明本文方法具有较高的拟合精度。与高斯分解相对比,本文由偏正态模型分解结果得到的地物相对高程信息更符合实际情况,证明了本文方法的有效性与优越性。在未来的工作中,针对特定全波 LiDAR 波形数据的理想模型的选取以及波形拟合精度的提高,还需进一步地研究。

## 参考文献:

- [1] 刘志青,李鹏程,陈小卫,等. 基于信息向量机的机载激光雷达点云数据分类 [J]. 光学精密工程, 2016, 24(1): 210-219.
- LIU ZH Q, LI P CH, CHEN X W, *et al.*. Classification of airborne LiDAR point cloud data based on information vector machine [J]. *Opt. Precision Eng.*, 2016, 24(1): 210-219. (in Chinese)
- [2] WAGNER W, ULLRICH A, DUCIC V, *et al.*. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2006, 60(2): 100-112.
- [3] 赖旭东,秦楠楠,韩晓爽,等. 一种迭代的小光斑 LiDAR 波形分解方法 [J]. 红外与毫米波学报, 2013, 32(4): 319-324.
- LAI X D, QIN N N, HAN X SH, *et al.*. Iterative decomposition method for small foot-print LiDAR waveform [J]. *Journal of Infrared and Millimeter Waves*, 2013, 32(4): 319-324. (in Chinese)
- [4] 周静平,张爱武,王书氏. 机载小光斑全波形 LiDAR 数据处理及应用 [J]. 测绘通报, 2013(1): 10-13.
- ZHOU J P, ZHANG A W, WANG SH M. Processing and application of small-footprinting full-waveform airborne LiDAR data [J]. *Bulletin of Surveying and Mapping*, 2013(1): 10-13. (in Chinese)
- [5] CHAUVE A, MALLETT C, BRETAR F, *et al.*. Processing full-waveform Lidar data: modelling raw signals[C]. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, ISPRS, 2007: 102-107.
- [6] 刘峰,谭畅. 全波形 LiDAR 数据分解方法的研究 [J]. 中南林业科技大学学报, 2010, 30(8): 148-154.
- LIU F, TAN CH. Study of decomposition of full waveform airborne laser scanner data [J]. *Journal of Central South University of Forestry & Technology*, 2010, 30(8): 148-154. (in Chinese)
- [7] 王素元,马洪超,王杰栋,等. 基于分组 LM 算法的全波形 LiDAR 高斯分解 [J]. 测绘与空间地理信息, 2016, 39(7): 144-147.
- WANG S Y, MA H CH, WANG J D, *et al.*. Gaussian decomposition of full-waveform LiDAR based on grouping LM algorithm [J]. *Geomatics & Spatial Information Technology*, 2016, 39(7): 144-147. (in Chinese)
- [8] PERSSONÅ, SÖDERMAN U, TÖPEL J, *et al.*. Visualization and analysis of full-waveform airborne laser scanner data[C]. *Workshop on "Laser Scanning 2005"*, ISPRS, 2005: 103-108.
- [9] 马洪超,李奇. 改进的 EM 模型及其在激光雷达全波形数据分解中的应用 [J]. 遥感学报, 2009, 13(1): 35-41.

- Ma H CH, LI Q. Modified EM algorithm and its application to the decomposition of laser scanning waveform data [J]. *Journal of Remote Sensing*, 2009, 13(1): 35-41. (in Chinese)
- [10] 王蕊, 邢艳秋, 孙小添, 等. 机载大光斑激光雷达数据估测森林结构参数研究进展 [J]. *遥感信息*, 2015, 30(3): 3-9, 23.  
WANG R, XING Y Q, SUN X T, *et al.*. Research progress on airborne large footprint LiDAR data in estimation of forest structure parameters [J]. *Remote Sensing Information*, 2015, 30(3): 3-9, 23. (in Chinese)
- [11] 庞勇. 星载干涉雷达和激光雷达数据森林参数反演 [D]. 北京: 中国科学院遥感应用研究所, 2005.  
PANG Y. *Forest Parameters Inversion Using Spaceborne in SAR and Lidar Technology* [D]. Beijing: Institute of Remote Sensing Application Chinese Academy of Science, 2005. (in Chinese)
- [12] 庞勇, 李增元, MICHAEL L, 等. 地形对大光斑激光雷达森林回波影响研究 [J]. *林业科学研究*, 2007, 20(4): 464-468.
- PANG Y, LI Z Y, MICHAEL L, *et al.*. Effects of terrain on the large footprint lidar waveform of forests [J]. *Forest Research*, 2007, 20(4): 464-468. (in Chinese)
- [13] 李玉, 徐艳, 赵雪梅, 等. 利用高斯混合模型的多光谱图像模糊聚类分割 [J]. *光学精密工程*, 2017, 25(2): 509-518.  
LI Y, XU Y, ZHAO X M, *et al.*. Multispectral image segmentation by fuzzy clustering algorithm used Gaussian mixture model [J]. *Opt. Precision Eng.*, 2017, 25(2): 509-518. (in Chinese)
- [14] 赵泉华, 李红莹, 李玉. 全波形 LiDAR 数据分解的可变量高斯混合模型及 RJMCMC 算法 [J]. *测绘学报*, 2015, 44(12): 1367-1377.  
ZHAO Q H, LI H Y, LI Y. Gaussian mixture model with variable components for full waveform LiDAR data decomposition and RJMCMC algorithm [J]. *Acta Geodaetica et Cartographica Sinica*, 2015, 44(12): 1367-1377. (in Chinese)

#### 作者简介:



赵泉华(1978—),女,河北承德人,博士,教授。2001年于河北理工大学获得学士学位,2004年、2009年于辽宁工程技术大学分别获得硕士、博士学位,主要研究方向为随机几何、空间统计学、模糊集理论等在遥感图像建模、解译及其在海洋环境遥感中的应用。E-mail: zhaquanhua@lntu.edu.cn



陈为多(1993—),女,辽宁阜新人,硕士研究生,2016年于辽宁工程技术大学获得学士学位,主要研究方向为全波 LiDAR 数据波形分解、全波 LiDAR 数据在地物提取方面的应用。E-mail: 644746512@qq.com