Research Article

A Denoising Method for LiDAR Full-Waveform Data

Xudong Lai^{1,2} and Min Zheng¹

¹*School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China*

²*Collaborative Innovation Center for Geospatial Technology, Wuhan 430072, China*

Correspondence should be addressed to Min Zheng; zhengmin1030@whu.edu.cn

Received 28 February 2015; Accepted 1 July 2015

Academic Editor: Ralph B. Dinwiddie

Copyright © 2015 X. Lai and M. Zheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Decomposition of LiDAR full-waveform data can not only enhance the density and positioning accuracy of a point cloud, but also provide other useful parameters, such as pulse width, peak amplitude, and peak position which are important information for subsequent processing. Full-waveform data usually contain some random noises. Traditional filtering algorithms always cause distortion in the waveform. λ/μ filtering algorithm is based on Mean Shift method. It can smooth the signal iteratively and will not cause any distortion in the waveform. In this paper, an improved λ/μ filtering algorithm is proposed, and several experiments on both simulated waveform data and real waveform data are implemented to prove the effectiveness of the proposed algorithm.

1. Introduction

The traditional airborne LiDAR (Light Detection and Ranging) systems only provide discrete multiple returns. They ignore the inherent information of the objects, such as structure and reflection characteristics. Furthermore, users cannot determine the errors caused by inadequate acquisition methods, because the acquisition methods are regarded as commercial secrets and are not public [1]. Since the Austrian company RIEGL first provides small-footprint scanners with capacity of digitizing entire waveform in commercial LiDAR system in 2004, more and more commercial LiDAR providers provide full-waveform data. Full-waveform scanning systems can sample the continuous laser signal by using a very small sampling interval and record it and not just record discrete return pulses [2]. The advantage is that users can process the full-waveform data and obtain corresponding information according to their applications.

Because of background light, detector noise, and noise of the digitization of waveform, there always exist messy data which do not belong to the laser signal, and the laser signals may be even mixed by the messy data [3]. Thus, properly removing the noise is important for subsequent processing of waveform data. The removal of random noise is a common problem in digital signal processing, especially in the field of

image processing. Traditional image filtering algorithms may lead significant distortion to the waveform, such as mean filtering and Gaussian filtering. And the filtering results directly influence the accuracy and reliability of waveform parameter calculation. Therefore, research on filtering method which can effectively keep the characteristics of LiDAR waveform data has important theoretical significance and practical value.

The noise of laser system includes additive noise and multiplicative noise, and both of these noises will have negative influence on applications of laser scanning data. Therefore, an important task in applications of laser scanning data is to find dominant noises and develop a method to remove the noises. For gas laser systems, additive noise plays a leading role, while the influence of multiplicative noise is negligible [4].

Signal denoising methods can be divided into two categories: frequency domain processing and spatial domain processing. Spatial domain denoising methods directly apply space transformation to the signals. It is usually a neighborhood operation, which means the value of a sampling point in the output signals is calculated by a certain algorithm by using its neighbor's value from the input signals. The algorithm can be average filtering, Gaussian filtering, and so forth. Frequency domain denoising methods first transform

the signal into frequency domain before the filtering and then transform the signal back to space domain by inverse transformation after the filtering, such as Fourier transform and wavelet transform.

The distribution of noises in LiDAR signals is very complicated, and only using the traditional digital filters that smooth the signals just by selecting a cut-off frequency is not efficient. In practice, the shape of waveform which is smoothed by the discrete wavelet transform is similar to wavelet function and contains serious distortions [5]. Fang and Huang [6] removed noises in signals by wavelet coefficients which used the nonlinear soft threshold technology and kept the feature of signals at the same time. However, the calculation quantity of wavelet multiscale decomposition reconstruction is too large to be used for smoothing thousands of pieces of LiDAR full-waveform data. Thus, its performance still needs to be further improved to solve practical problems.

As for the space domain denoising, except for traditional image processing algorithms, Sun et al. [7] regarded LiDAR signal denoising as a function regression problem and used the least squares support vector machine (SVM) to remove the noise in LiDAR data. However, if a better denoising result is needed, the prior knowledge of LiDAR signals should be added to the training of the support vector machine. The drawback of this method is that if the distribution of LiDAR waveform is not the sum of Gaussian distribution, the denoising result will not be ideal. However, LiDAR waveform signal is not strict Gaussian distribution in practice.

The traditional image spatial domain filtering algorithms such as mean filtering and Gaussian filtering do not need a priori knowledge of the waveform data, and the algorithms are simple and suitable for rapid processing of a large number of waveform data. But using these algorithms for denoising may cause distortion of the waveform, such as the shrinkage of the peak amplitude and the increase of pulse width [8]. Yang and Huang [9] adopted a Gaussian filter to remove random noise in waveform data and pointed out the drawbacks of traditional denoising algorithms.

The organization of this paper is as follows: the second part mainly introduces the improved λ/μ filtering algorithm; the third part presents experiments on both simulated waveform data and real waveform data and compared the experimental results; finally a summary of the whole paper is given.

2. Improved λ/μ Filtering Algorithm

λ/μ filtering algorithm is proposed by Taubin, which is based on Mean Shift method [10]. This algorithm is a low-pass filter defined on a Gaussian kernel function. It adds a factor which can suppress the shrinkage of peak amplitude to Gaussian function and effectively overcomes the drawback of Gaussian filtering that leads to waveform edge shrinkage, thus avoiding the waveform distortion effectively.

2.1. Traditional λ/μ Filtering Algorithm. A one-dimensional signal x can be represented as a column vector $(x_1, \dots, x_n)^T$,

and the simplest form of Gaussian filtering method can be described by

$$x'_i = x_i + \lambda \Delta x_i, \quad (1)$$

where $\Delta x_i = (1/2)(x_{i-1} - x_i) + (1/2)(x_{i+1} - x_i)$, λ is a scale factor, and $0 < \lambda < 1$. The formula can be written in matrix form as

$$x' = (I - \lambda K) x, \quad (2)$$

where K is the matrix as

$$K = \frac{1}{2} \begin{pmatrix} 2 & -1 & & -1 \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \\ -1 & & & -1 & 2 \end{pmatrix}. \quad (3)$$

The Gaussian filtering may lead shrinkage to the edge of the signal, and the Gaussian kernel function cannot be the low-pass filter kernel. To solve this problem, a nonshrinking function $f(K)$ of the matrix K is chosen to replace the matrix $I - \lambda K$, so

$$x' = f(K) x \quad (4)$$

and this process can be iterated N times:

$$x^N = f(K)^N. \quad (5)$$

Since K is symmetric, it has real eigenvalues and eigenvectors. If the real eigenvalues of K are $0 \leq k_1 \leq k_2 \leq \dots \leq k_n \leq m$ and the corresponding eigenvectors are $u_1, u_2, \dots, u_n$, then formula (4) can be described as

$$x' = f(K) x = \sum_{i=1}^n \xi_i f(k_i) u_i, \quad (6)$$

where f is filtering kernel function. If $k \in [0, m]$, after iterating for N times, in low frequency part $f(k_i)^N \approx 1$, while in high frequency part $f(k_i)^N \approx 0$. Thus the choice of kernel function is shown as

$$f(k) = (1 - \lambda k)(1 - \mu k), \quad (7)$$

where μ is a negative scale factor and $\mu < -\lambda$. It means that after smoothing by formula (1) with positive scale factor another calculation step was taken

$$x'_i = x_i + \mu \Delta x_i. \quad (8)$$

Since $f(0) = 1$, and $\lambda + \mu < 0$, chose a marginal value k_{PB} fulfill $f(k_{PB}) = 1$, λ and μ meet the condition:

$$k_{PB} = \frac{1}{\lambda} + \frac{1}{\mu} > 0. \quad (9)$$

Smoothing by (1) will lead distortion to signal because all the calculations are additive operation; if the calculation of (8) is added and calculated as the two formulas alternately, the drawback of Gaussian filtering can be overcome. This is the principal of λ/μ filtering algorithm.

2.2. Improved λ/μ Filtering Algorithm. The traditional λ/μ algorithm is based on Gaussian filtering and modified the kernel function. It restrains the shrinkage caused by Gaussian filtering through the method of the alternate addition and subtraction calculation. However, since $\mu < -\lambda$, the degree of addition calculation and subtraction calculation is different. As a consequence, the result of denoising is affected. In this paper, the algorithm is improved. The traditional algorithm will operate for several times. When it is the odd time, use the following formula:

$$\begin{aligned} x'_i &= x_i + \lambda \Delta x_i & \text{when } i \text{ is odd} \\ x'_i &= x_i + \mu \Delta x_i & \text{when } i \text{ is even;} \end{aligned} \quad (10)$$

when it is the even time, use the following formula:

$$\begin{aligned} x'_i &= x_i + \lambda \Delta x_i & \text{when } i \text{ is even} \\ x'_i &= x_i + \mu \Delta x_i & \text{when } i \text{ is odd.} \end{aligned} \quad (11)$$

Each processing of the traditional algorithm is weighted correction on a single direction, and the directions of correction between adjacent points are different. Thus, processing for several times may decrease the difference between the point and the mean value of its neighbors; the difference between two adjacent points may be significant. As a result, the elimination of noise is not effective enough. The improved algorithm uses the traditional algorithm alternately; namely, the directions of different correction of the same point are different. The iterative processing can offset the distortion caused by the algorithm, and the denoising result is further improved.

3. Denoising Experiment

In order to evaluate the effectiveness of each filtering algorithm, this paper uses both simulated waveform data and real waveform data for the experiments and comparison.

3.1. Experiment on Simulated Data. LiDAR full-waveform data is composed of n equidistant discrete points, where n is the sampling number, and the distance between two adjacent points is sampling interval. Different data acquisition systems have different sampling numbers and sampling intervals. In general, sampling number could be 64, 128, or 256. LiDAR waveform data is similar to the composition of Gaussian distribution [11], and it has been confirmed from two aspects of mathematics and physics by Li and Ma [12]. The sampling number of simulated waveform data is 128, and it is composed of three Gaussian distributions. Their peak amplitude, peak position, and pulse width are (84.29, 25.00, 4.80), (28.99, 19.00, 2.40), and (17.16, 33.00, 2.40), and the parameters are acquired through the decomposition of real LiDAR full-waveform data. The simulated data without noise is shown in Figure 1. The abscissa is sampling number, and y coordinate is amplitude (the same in all the figures).

The simulated waveform data only considers the additive noise and is simulated by adding salt and pepper noise to the ideal nonnoise Gaussian wave. Salt and pepper noise in

FIGURE 1: Simulated ideal nonnoise Gaussian wave.

FIGURE 2: Simulated Gaussian wave with noise.

the image refers to the black and white noise, including salt noise and pepper noise. The former is high gray noise and the latter is low gray noise. In waveform data, salt and pepper noise refers to plus or minus a suitable value in some of the sampling points of the ideal nonnoise waveform and makes the ideal waveform containing noise. In this paper, 35 noises are created by using Matlab and the signal-to-noise ratio is 6 watts. Then the noises are added to the ideal nonnoise Gaussian wave randomly. In practice, there is backnoise in real waveform data; however, the removal of backnoise will not be discussed in this paper. Thus backnoise with constant value of 10 is added to the simulated data. Waveform with noise is shown in Figure 2.

Figure 3 shows the denoising results of salt and pepper noise by mean filtering, Gaussian filtering, traditional λ/μ algorithm filtering, and improved λ/μ algorithm filtering when the size of processing window is 3, and the parameters of λ/μ algorithm are selected as follows: $k_{PB} = 0.1$, $\lambda = 0.6307$, and $\mu = -0.6372$.

As shown in Figure 3, when the size of processing window is 3, all the algorithms cannot achieve a fine denoising result, and the noise is still obvious. Moreover, when processing window is 7, the data will be overfiltered and will lose some detailed information. In order to achieve more obvious denoising results and retain the details of waveform data, the size of processing window is enlarged to 5, and the denoising

FIGURE 3: Results of denoising by different algorithms when the size of processing window is 3 (the black point is simulated waveform, the red curve is denoised waveform, and the blue curve is the ideal nonnoise waveform).

results are shown in Figure 4. In addition, wavelet denoising is brought in as a representative of frequency domain denoising methods to be compared with the proposed method. This paper uses soft threshold wavelet denoising and the threshold is calculated by the formula $\sigma\sqrt{2\ln n}$, where σ is the standard deviation of noise and n is the length of the waveform [13].

Figure 4 shows that the denoising results are much better after enlarging the size of processing window. However, mean filtering and Gaussian filtering lead obvious distortion to the waveform, and the peak becomes smaller and width becomes larger seriously. Wavelet denoising reaches a good result except the distortion of waveform. Traditional λ/μ filtering algorithm can smooth waveform and only cause very slight distortion, because λ/μ filtering algorithm added a factor which can suppress the shrinkage of the peak based on Gaussian function and overcome the drawback of Gaussian filtering that led to waveform edge shrinkage. Thus, it can smooth waveform effectively and keep the shape of waveform at the same time. As shown in Figure 5, the waveform smoothed by improved λ/μ algorithm is nearer to the original waveform than that smoothed by traditional λ/μ algorithm and the value difference is up to 1. The square deviation is imported to evaluate the effectiveness of these methods.

The calculation of peak signal-to-noise ratio is shown as the following formula:

$$\text{square deviation} = \frac{\sum_{i=1}^{128} (x_i - x'_i)^2}{128}, \quad (12)$$

where x_i is the i th sampling point of original waveform and x'_i is the i th sampling point of filtered waveform. This parameter shows the difference between original waveform and filtered waveform, and the smaller the parameter is, the filtered waveform is more similar with the original waveform. Table 1 is the square deviation of different methods and it shows that denoising result of improved λ/μ filtering algorithm is better than traditional λ/μ filtering algorithm, for the denoising result of improved algorithm is more close to the ideal nonnoise waveform. Table 2 compares the parameters of waveform processed by different algorithms and shows the advantage of improved λ/μ filtering algorithm in maintaining the shape of waveform.

3.2. Experiment on Real Data. The real waveform data in this paper is collected by the Leica ALS60 system and saved as LAS1.3 format. The surveying area is urban, and it contains building, tree, and ground points. The sampling

FIGURE 4: Results of denoising by different algorithms when the size of processing window is 5 (the black point is simulated waveform, the red curve is denoised waveform, and the blue curve is the ideal nonnoise waveform).

number of each waveform is 128, and the sampling interval is 1 nanosecond. The experiments on simulated data show that when the size of processing window is 3, no matter which filtering algorithm is used, the result of denoising is not remarkable. When it comes to real data, the size of processing window is directly chosen as 5. Figure 6 shows the denoising results of real waveform data by wavelet denoising, mean filtering, Gaussian filtering, traditional λ/μ filtering algorithm, and improved λ/μ filtering algorithm when the size of processing window is 5.

Figure 6 shows that wavelet denoising, mean filtering, and Gaussian filtering cause obvious waveform distortion and the peak decreases seriously. The difference between the results of traditional and improved λ/μ filtering algorithm is not significant in figures, so the peak signal-to-noise ratio (PSNR) is imported to evaluate the effectiveness of these two methods. The calculation of peak signal-to-noise ratio is shown as formula (13). In the formula, the max is the maximum value in the waveform, and x_i is the value of the i th point in the waveform, and x_i^* is the value of the i th point

FIGURE 5: The details of Figure 4 (the black point is simulated waveform, the red curve is denoised waveform, and the blue curve is the ideal nonnoise waveform).

TABLE 1: The square deviation of waveform denoising.

Algorithm	Wavelet denoising	Mean filtering	Gaussian filtering	Traditional λ/μ algorithm filtering	Improved λ/μ algorithm filtering
Square deviation	1.86	2.69	6.29	1.26	0.89

TABLE 2: The comparison of parameters of simulated waveform.

Algorithm	Original waveform	Wavelet denoising	Mean filtering	Gaussian filtering	Traditional λ/μ algorithm filtering	Improved λ/μ algorithm filtering
Peak position	25	26	25	25	25	25
Peak amplitude	91.45	88.74	86.78	84.59	90.03	90.89

TABLE 3: The PSNR of real waveform denoising.

Algorithm	Wavelet denoising	Mean filtering	Gaussian filtering	Traditional λ/μ algorithm filtering	Improved λ/μ algorithm filtering
PSNR	37.28	36.76	32.26	41.17	44.46

in the denoised waveform. The PSNR of denoised waveform is shown in Table 3, and it shows that the PSNR of denoising result of improved λ/μ filtering algorithm is larger:

$$\text{PSNR} = 10 \times \log \frac{\sum_{i=1}^{128} \max^2}{\sum_{i=1}^{128} (x_i^* - x_i)^2}. \quad (13)$$

As shown in Figure 7, the shrinkage of waveform peak smoothed by improved λ/μ algorithm is not as serious as that smoothed by traditional λ/μ algorithm. In addition, the parameters of waveforms are compared in Table 4. It is obvious that filtering by improved λ/μ algorithm can maintain the peak better. Because the sampling number of the

waveform data is only 128, the peak position differences are not obvious.

4. Conclusion

In data processing, noise must be considered. Only by removing noise effectively fine results can be achieved in further process. This paper introduced the common methods of random noise removal and improved the traditional λ/μ filtering algorithm. The effectiveness of frequency domain denoising, traditional image processing algorithms, λ/μ filtering algorithm, and improved λ/μ filtering algorithm is compared by experiments on both simulated data and real data. Based on the experiments, this paper pointed out that

FIGURE 6: Results of denoising by different algorithms when the size of processing window is 5 (the point is real waveform, and the curve is denoised waveform).

TABLE 4: The comparison of parameters of real waveform.

Algorithm	Original waveform	Wavelet denoising	Mean filtering	Gaussian filtering	Traditional λ/μ algorithm filtering	Improved λ/μ algorithm filtering
Peak position	25	24	25	24	25	25
Peak amplitude	108	103.02	103.20	102.85	106.96	107.68

FIGURE 7: Details of Figure 6 (the point is real waveform, and the curve is denoised waveform).

denoising by traditional image processing algorithms and wavelet denoising will cause distortion to the LiDAR waveform and proved the superiority of λ/μ filtering algorithm. λ/μ filtering algorithm can remove random noise effectively and keep the shape of the waveform at the same time, and the improved λ/μ filtering algorithm can further improve the signal-to-noise ratio, which is useful for the further processing of waveform data.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. Reitberger, P. Krzystek, and U. Stilla, "Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees," *International Journal of Remote Sensing*, vol. 29, no. 5, pp. 1407–1431, 2008.
- [2] X. D. Lai, *Airborne LiDAR Basic Principle and Application*, Publishing House of Electronics Industry, Beijing, China, 2010.
- [3] X. Li, J. B. Gong, and Z. W. Zhou, "Laser imaging radar waveform digitalize technology," *Infrared and Laser Engineering*, no. 1, pp. 474–477, 2007.
- [4] J. Hu, "Noise in laser," *Journal of Chifeng College*, vol. 26, no. 5, pp. 112–113, 2010.
- [5] H. Y. Zhang, G. H. Fan, T. H. Zhang, and Y. H. Zheng, "Wavelet denoising study of laser radar waveform signal," *Research and Development*, vol. 5, no. 31, pp. 52–58, 2012.
- [6] H.-T. Fang and D.-S. Huang, "Noise reduction in lidar signal based on discrete wavelet transform," *Optics Communications*, vol. 233, no. 1–3, pp. 67–76, 2004.
- [7] B.-Y. Sun, D.-S. Huang, and H.-T. Fang, "Lidar signal denoising using least-squares support vector machine," *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 101–104, 2005.
- [8] L. L. Liu, *Decomposition of Airborne LiDAR Full-Waveform Data Based on LM Method*, Wuhan University, Wuhan, China, 2011.
- [9] G. Yang and C. M. Huang, "Decomposition algorithm of laser altimeter," *Chinese Journal of Space Science*, vol. 25, no. 2, pp. 125–131, 2005.
- [10] G. Taubin, "Signal processing approach to fair surface design," in *Proceedings of the 22nd Annual ACM Conference on Computer Graphics and Interactive Techniques*, pp. 351–358, ACM, New York, NY, USA, August 1995.
- [11] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 100–112, 2006.
- [12] Q. Li and H.-C. Ma, "The study of point-cloud production method based on waveform laser scanner data," *Acta Geodaetica et Cartographica Sinica*, vol. 37, no. 3, pp. 349–354, 2008.
- [13] D. L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *Applied and Computational Harmonic Analysis*, vol. 1, no. 1, pp. 100–115, 1993.