



Article

True-Color Reconstruction Based on Hyperspectral LiDAR Echo Energy

Tengfeng Wang ¹ , Xiaoxia Wan ^{1,2,*}, Bowen Chen ^{3,4}  and Shuo Shi ⁴

¹ School of Printing and Packaging, Wuhan University, Wuhan 430079, China; wangtengfeng@whu.edu.cn

² Hubei Province Engineering Technical Center for Digitization and Virtual Reproduction of Color Information of Cultural Relics, Wuhan 430079, China

³ Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China; chenbowen1204@whu.edu.cn

⁴ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; shishuo@whu.edu.cn

* Correspondence: wan@whu.edu.cn

Abstract: With the development of remote sensing technology, the simultaneous acquisition of 3D point cloud and color information has become the constant goal for scientific research and commercial applications in this field. However, since radar echo data in practice refer to the value of the spectral channel and its corresponding energy, it is still impossible to obtain accurate tristimulus values of the point through color integral calculation after traditional normalization and multispectral correction. Furthermore, the reflectance of the target, the laser transmission power and other factors lead to the problems of no echo energy or weak echo energy in some bands of the visible spectrum, which further leads to large chromatic difference compared to the color calculated from the spectral reflectance of standard color card. In response to these problems, the hyperbolic tangent spectrum correction model with parameters is proposed for the spectrum correction of the acquired hyperspectral LiDAR in the 470–700 nm band. In addition, the improved gradient boosting decision tree sequence prediction algorithm is proposed for the reconstruction of missing spectrum in the 400–470 nm band where the echo energy is weak and missing. Experimental results show that there is relatively small chromatic difference between the obtained spectral information after correction and reconstruction and the spectrum of standard color card, achieving the purpose of true color reconstruction.

Keywords: true-color reconstruction; multispectral correction; spectrum reconstruction; hyperspectral LiDAR



Citation: Wang, T.; Wan, X.; Chen, B.; Shi, S. True-Color Reconstruction Based on Hyperspectral LiDAR Echo Energy. *Remote Sens.* **2021**, *13*, 2854. <https://doi.org/10.3390/rs13152854>

Academic Editor: P. Takis Mathiopoulos

Received: 27 May 2021

Accepted: 19 July 2021

Published: 21 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



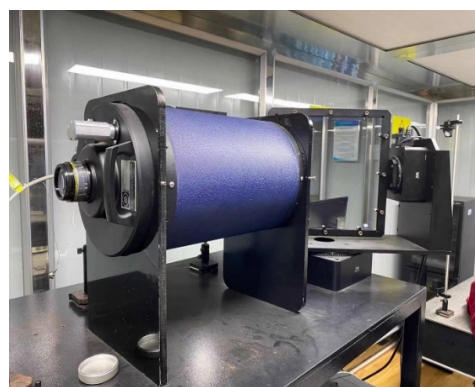
Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

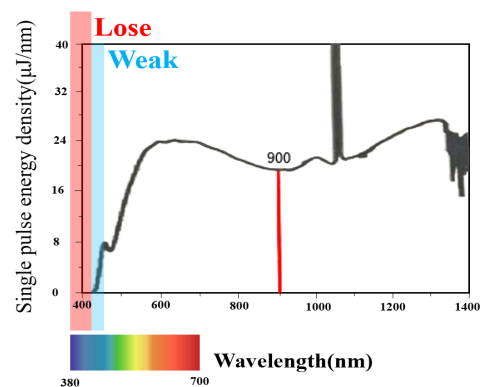
With the emergence of projects such as “Realistic 3D” and “Smart City”, there are increasingly urgent needs for target true color reproduction in the fields of urban planning, landform restoration, and scene reconstruction [1–7]. In order to reproduce the 3D urban real scene or reflect the real geographic pattern and information, the structure and color of the target or landform must be accurately restored [8]. In the visible range, spectral reflectance is the fingerprint of target surface color, which characterizes the ratio of the radiant exitance on object surface to the radiant incidence at each waveband, and reflects the object's absorption and reflection characteristics of radiant energy [9–11]. According to the colorimetry principle, the tristimulus values $\{X, Y, Z\}$ [12] of the object color are obtained by integrating the relative spectral power distribution of illuminant $I(\lambda)$, spectra of object surface $r(\lambda)$, standard colorimetric observer matching functions $\{\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)\}$, adjustment factor K , and wavelength sampling interval $d(\lambda)$, as shown in Equation (1).

$$\begin{aligned}
 X &= K \int l(\lambda) r(\lambda) \bar{x}(\lambda) d\lambda; \\
 Y &= K \int l(\lambda) r(\lambda) \bar{y}(\lambda) d\lambda; \\
 Z &= K \int l(\lambda) r(\lambda) \bar{z}(\lambda) d\lambda;
 \end{aligned} \tag{1}$$

As shown in Figure 1a, the new hyperspectral LiDAR [13–19] not only retains the three-dimensional space detection capability of laser radar, but also increases the number of laser detection bands. By combining the advantages of active detection technology and passive detection technology, the integrated acquisition of spatial 3D spectral information on the same device is realized, which increases the possibility of conducting 3D reconstruction and true-color restoration simultaneously [20]. However, since the laser radar echo data are the energy values corresponding to the spectral channel, there is relatively large spectral difference between the spectral response function calculated by the traditional normalization method and the spectrum correction with the standard spectrum provided by the X-rite ColorChecker Classic in practical applications, which leads to the colour difference in the color reconstruction. Moreover, the distribution of laser transmission power also causes some problems. As shown in Figure 1b, within the visible light range of 400–700 nm, there is no echo energy when the band is 400–430 nm, and the echo energy is weak when the band is 430–470 nm, which leads to the problem that the target cannot be correctly reconstructed with true color. Therefore, conduct normalization and spectrum correction on laser radar echo energy data, and perform spectrum reconstruction on the missing spectral information in the band of 400–470 nm have become the key to the true-color restoration of the target.



(a) Multispectral LiDAR



(b) Laser radar divergence power and band

Figure 1. Laser radar and power diagram.

In terms of true-color reconstruction of hyperspectral LiDAR, Wang et al. [21] and Chen et al. [22] found that the most suitable wavelength combinations for RGB true-color reconstruction are 466 nm, 546 nm and 626 nm through the analysis of the true-color synthesis results of different wavelengths. Therefore, they selected three wide bands near the corresponding bands, especially the relatively wide blue band, to increase the energy of the RGB channel and obtain a higher signal-to-noise ratio. This method can achieve the color reconstruction of the target to a certain extent. On the other hand, due to the metamerism problem and the failure to consider the standard illuminant and standard observer, the color obtained by reconstruction will have color difference after changing the illuminant.

Normalization and spectrum correction, as the key in hyperspectral laser radar energy data processing, play a decisive role in the accuracy of the subsequent color integral calculation results of multispectral data. Traditional spectrum correction algorithms include first

derivative spectrophotometry, centroid algorithm local similarity matching algorithm and various improvement methods [23]. Among them, first derivative spectrophotometry [24] and local similarity matching algorithm [25] have the advantages of simple operation and high real-time, as well as the disadvantage of being easily affected by noise. The centroid algorithm [26,27] is sensitive to the spectrum waveform asymmetry, which will cause the corrected multispectral image to generate random direction deviation. After simulation correction of radar echo energy through the above algorithms, the color obtained by spectral response curve integral calculation and standard color block calculation still has great color difference, which can cause the traditional spectrum correction method to fail to perform true-color reconstruction on the target.

Regarding the reconstruction of missing spectrum, the missing band fixed by laser radar is between 400–470 nm, and the information in the 470–700 nm band is relatively intact. At the same time, there is a strong correlation between the multispectral spectrum segments. Therefore, it is feasible to restore the missing 400–470 nm band by using 470–700 nm. Adam et al. [28] sampled the low-quality spectrum segments in multispectral data and the adjacent high-quality spectrum segments respectively, and rearranged them to form a 3D data cube. Then, wavelet decomposition and sparse transformation were used to restore the data cube. Zhong et al. [29] firstly established the conditional random fields model between multiple spectrums to select the unusable part of the spectrum and the spectrum segment with high signal-to-noise ratio. Then, the texture and statistical information related to the spectrum segment with high signal-to-noise ratio was used to restore the continuous missing spectrum segment. Yin et al. [30] believed that the spectral information of multispectral images was highly redundant. They discussed the sparsity of the spectrum by setting up the general Gaussian model. In addition, they conducted the spectrum-dimensional reconstruction based on the compressed sensing theory, which was superior to the linear interpolation in the recovery of missing spectrum. Liu et al. [31] established a bi-inverted Gaussian fitting model. As atmospheric water vapor absorbed the reconstruction vegetation and resulted in the unusable spectrum, the diagnostic spectrum features of vegetation absorption peak were extracted for the inversion of water content. It should be noted that the above model-based reconstruction methods for missing spectrum only have good effects for plants within certain wavelength ranges or specific types of geographic objects, and their versatility in true-color reconstruction is significantly limited.

In order to solve the problem of the large spectral difference between the normalized echo energy data based on traditional methods and the standard spectrum provided by X-rite ColorChecker Classic, this article proposes a Hyperbolic Tangent Spectrum Correction Model with Parameters. The echo energy processed by this model has a relatively small spectral difference compared with the X-rite Eye One Pro spectrophotometer at 10 nm intervals from 400 to 700 nm.

In order to solve the problem of serious color difference in true-color restoration due to the missing of 400–470 nm band spectrum, a series forecasting algorithm based on gradient boosting decision tree is proposed in this article. The spectrum data provided by X-rite ColorChecker Classic is used as the radar control experiment data, and the multispectral data set [32] provided by ICVL (interdisciplinary computational vision laboratory) is used as the prior data and test data. Through cyclic learning, the spectrum data are divided into 31 parts, and each feature is modeled separately. Through the locally weighted linear regression algorithm, the missing 400–470 nm band information is cyclically and accurately reconstructed, so as to solve the problem of great color difference in the true-color reconstruction result.

Through the spectral reconstruction by hyperbolic tangent normalization and correction model with parameters and improved gradient boosting decision tree sequence prediction algorithm, the relative energy data and coordinate point cloud data of the radar echo can be directly processed into the color information and position information of each point, thus achieving the integrated representation of echo energy-spectrum-color.

2. HSL System Description

All spatial point cloud and echo energy data in this article were obtained through the all-day active hyperspectral LiDAR prototype device. The laboratory HSL is mainly composed of five parts, including supercontinuum laser, optical receiver, light-splitting, signal reception and ranging components. First, high-nonlinearity photonic crystal fibers are used to generate and emit wide-band white lasers. Among them, the output power of the supercontinuum laser source is greater than 6 W, with the repetition rate of 0.1–1 MHz and the pulse duration of 4 ns. The spectral range of the laser is 400–2400 nm. In order to consider the data quality of backscattered signals, a high-precision two-dimensional scanning platform is adopted, and it is equipped with off-axis parabolic mirror of high reflectivity and Schmidt-Cassegrain telescope with focal length and diameter of 400 mm and 200 mm. The collected signal is guided by fiber to a grating spectrometer. In order to accurately project the spectral signals onto the detectors, we adopted a 150 g/mm blazed grating with high spectroscopic efficiency. The 32-element photosensitive photomultiplier tube array with a wavelength of 300–920 nm is selected as the received sensor. The output signal of the laser source and the backscattered signals of the target will be recorded into the APD detector respectively. By comparing the time interval between output signals and backscattered signals, time-of-flight can be used to calculate distance. The laser radar samples the multispectral data and point cloud data for each point on the X-rite ColorChecker Classic, as shown in Figure 2. The standard color card used was X-rite ColorChecker Classic, and the sampling method was the single point test of the color sample of color card. The final value was the average of three samplings at a single point. The collected standard diffuse reflection whiteboard channel energy values were used as the peak value of the traditional normalization method.

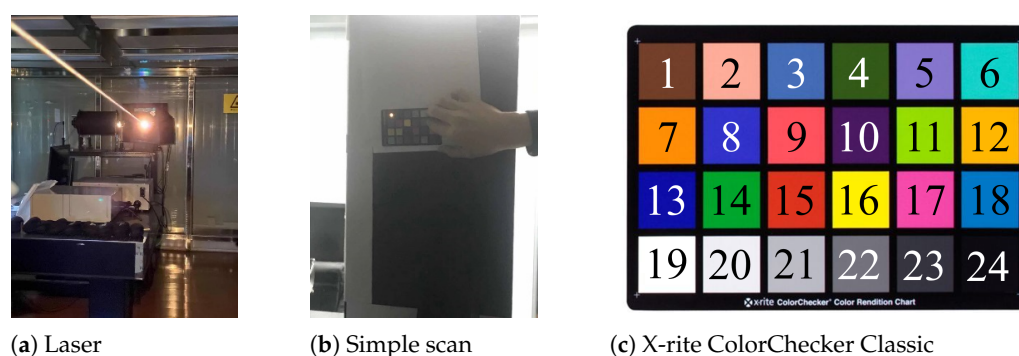


Figure 2. hyperspectral LiDAR data collection.

3. Hyperbolic Tangent Normalization and Correction Model with Parameters

After sampling the energy of X-rite ColorChecker Classic hyperspectral LiDAR, the corresponding energy distribution of each band is shown in Figure 3. By analyzing the corresponding relationship between energy and spectrum, the following findings are obtained. When the spectral band is between 400 nm and 430 nm, the echo energy obtained by the hyperspectral LiDAR is almost 0. As a result, the spectral information of this band is completely lost. When the spectral band is between 430–470 nm, the echo energy obtained by hyperspectral LiDAR is less than 5000. Therefore, it is a weaker echo energy range. When the spectral band is between 470–700 nm, the hyperspectral LiDAR has relatively normal echo energy data. Within the visible light range of 400–700 nm, the energy curve also corresponds to the divergence power of the laser radar, as shown in Figure 1b.

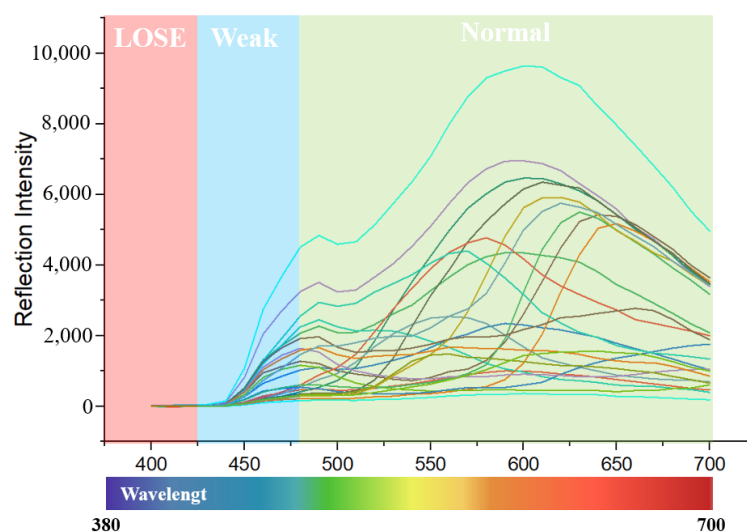


Figure 3. Laser radar channel and energy diagram.

Although the echo energy in the 470–700 nm range is normal, by measuring the standard diffuse reflection whiteboard, the spectral response curve of the normalized echo energy is significantly different from the standard spectral response curve provided by the X-rite ColorChecker Classic color card, as shown in Figure 4.

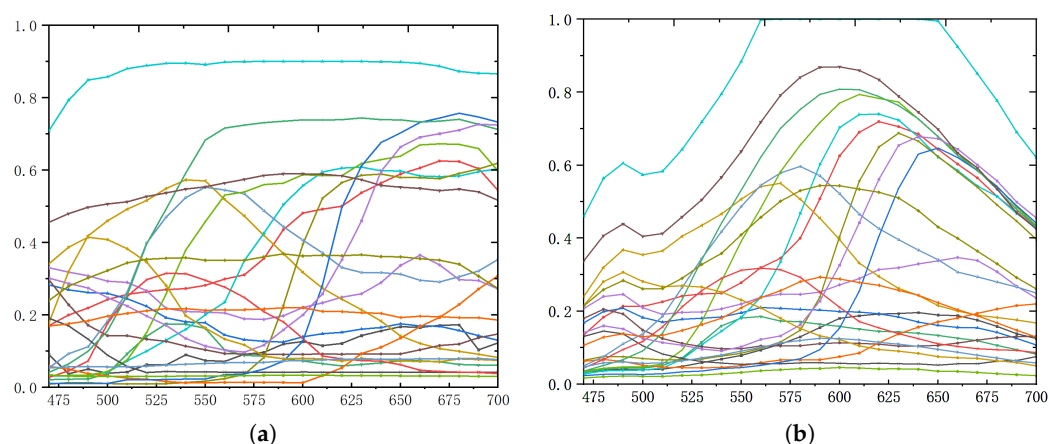


Figure 4. Normalized spectral response curve by whiteboard. (a) Spectral response curve provided by X-rite ColorChecker Classic color card, (b) Normalized spectral response curve measured by hyperspectral LiDAR.

The expected spectral response curve has the higher sensitivity and change rate when the echo energy is low, as well as the lower sensitivity and change rate when the reflected energy approaches the maximum value. In order to get the accurate spectrum normalization value, this article proposes a new normalization model. The min-max normalization of the energy function in the 470–700 nm band is shown in Equation (2):

$$f(x) = \frac{x - \text{MIN}}{\text{MAX} - \text{MIN}} \quad (2)$$

where, MIN is the minimum value among all the energy values collected in the 470–700 nm band, and MAX is the maximum value. By comparing the spectral response curves, we found that the two curves have the same trend but have larger fluctuations. The function fitting method can solve this problem well. When it is close to 1, it can be flattened, and the other parts retain their original sensitivity. Hyperbolic tangent function has this

property. As shown in Equation (3), the hyperbolic tangent spectrum correction model with parameters is used to correct the normalization value in the 470–700 nm band.

$$f(x, b) = \frac{\exp(b \cdot x^2) - \exp(-b \cdot x^2)}{\exp(b \cdot x^2) + \exp(-b \cdot x^2)} \quad (3)$$

In the calibration process, the mean square difference between the target spectrum and the source spectrum information $\exists b$ is calculated. By making $f(x, b) = 0$ and substituting the parameter b into the hyperbolic tangent calibration model, the normalized spectral response curve of 470–700 nm band is obtained.

4. Improved Spectrum Reconstruction of Gradient Boosting Decision Tree Series Forecasting

In essence, the gradient boosting decision tree is a kind of integrated learning method based on tree structure. The basic tree structure is the classification and regression tree, which is a locally weighted linear regression algorithm that divides the features into multiple parts and models each feature separately. Different from the traditional methods, the regression algorithm based on the tree structure is a parameter-based learning algorithm. Through the data training of model, the parameter value of each tree node is determined, without the need of updating in the subsequent prediction process.

Currently, Gradient Boosting Decision Trees [33–35] has many improved algorithms, but they are all for classification tasks. Zhang J., et al. [36] proposed Fisher score and Gradient Boosting Decision Trees algorithm, this paper combines Fisher algorithm with features trained in the gradient boosting decision tree for cancer risk identification. Zhang W., et al. [37] This paper is also an improved Gradient Boosting Decision Trees for the prediction model for air conditioning system loads. Chen Yiyi, et al. [38] The paper is a classification of data related to psychology by using gradient boosting decision trees. All of the above are improvements to classification problems using gradient boosting decision trees, which are not suitable for trend prediction tasks.

Firstly, the gradient boosting decision tree has been rewritten from the classification method to the sequence prediction algorithm. Therefore, the output value is no longer categories, but sequence prediction results. Secondly, if the correction function is not added, the subsequent value will become larger and larger, thus affecting the accuracy of prediction result.

It is assumed that the model is represented as $F(x; P)$, where P represents the parameter set, $P = \{p_0, p_1, p_2 \cdots p_n\}$, and $F(x; P)$ represents the prediction function x with P as the parameter. As each feature is modelled separately, the model is composed of multiple sub models, where β represents the weight of each sub model, and α represents the parameters in each sub model. Optimizing $F(x; P)$ can be abstracted to optimizing $\{\beta, \alpha\}$, i.e., optimizing parameter P . Considering that the optimized data set is normalized based on Equation (3), the expression for x in the form of Equation (3) is used to achieve consistency between the optimized result and the expected result.

$$\begin{aligned} F(x; P) &= F\left(\frac{\exp(b \cdot x^2) - \exp(-b \cdot x^2)}{\exp(b \cdot x^2) + \exp(-b \cdot x^2)}; \{\beta_m, \alpha_m\}_1^M\right) \\ &= \sum_{m=1}^M \beta_m h\left(\frac{\exp(b \cdot x^2) - \exp(-b \cdot x^2)}{\exp(b \cdot x^2) + \exp(-b \cdot x^2)}; \alpha_m\right) \end{aligned} \quad (4)$$

Equation (4) performs a hyperbolic optimization of the expression for x . As it approaches 1, the model has less variation. At it is less than 1, the model can maintain its original sensitivity. The parameter b is derived by Equation (3). The parameter $\{\beta, \alpha\}$ in the model is represented by Equation (5):

$$\{\beta_m, \alpha_m\}_1^M = \arg \min \sum_{i=1}^N L \left(y_i, \sum_{m=1}^M \beta_m h(x_i; \alpha_m) \right) \quad (5)$$

The purpose of this equation is to use N sample points (x_i, y_i) to calculate the loss function under the model $F(x; \beta, \alpha)$. The purpose of optimizing $\{\beta, \alpha\}$ is to minimize the value of the loss function. The model $F_m(x)$ is optimized in the direction where the loss function of the previously obtained model $F_{m-1}(x)$ declines fastest. Each point x_i yields $\overrightarrow{g_m(x_i)}$. Equation (6) is used to derive the complete direction of gradient descent.

$$\overrightarrow{g_m} = \{-g_m(x_i)\}_1^N \quad (6)$$

In order for $F_m(x)$ to be in the direction of $\overrightarrow{g_m(x_i)}$, least-square method is used to optimize the Equation (5) to get Equation (7):

$$\alpha_m = \arg \min \sum_{i=1}^N (-g_m(x_i) - \beta h(x_i; \alpha))^2 \quad (7)$$

On the basis of the derived α_m , β_m can be calculated, as shown in Equation (8):

$$\beta_m = \arg \min \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \beta h(x_i; \alpha_m)) \quad (8)$$

The final combined model equation is shown in Equation (9):

$$F_m(x) = F_{m-1} \left(\frac{\exp(b \cdot x^2) - \exp(-b \cdot x^2)}{\exp(b \cdot x^2) + \exp(-b \cdot x^2)} \right) + \beta_m h \left(\frac{\exp(b \cdot x^2) - \exp(-b \cdot x^2)}{\exp(b \cdot x^2) + \exp(-b \cdot x^2)}; \alpha_m \right) \quad (9)$$

where, the value of parameter b is the parameter value calculated by the normalization equation. The overall algorithm flow chart is shown in Figure 5, where the parameter P is the number of channels to be predicted, and M is the number of features.

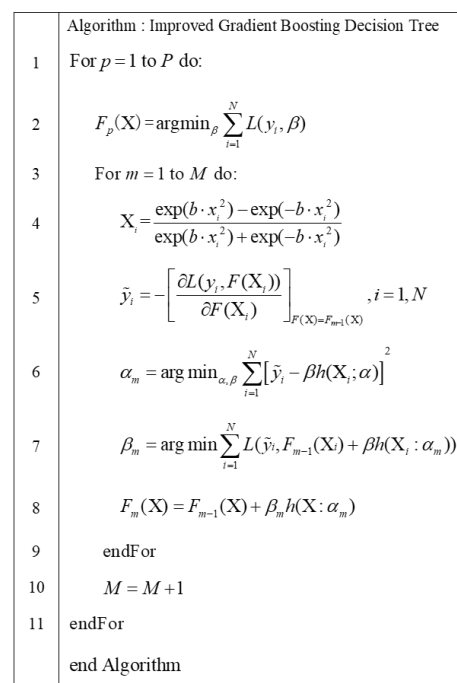


Figure 5. Algorithm flow chart.

In addition, there are many trend prediction algorithms, such as seq2seq and Random-Forest. However, Seq2Seq and RandomForest are suitable for NLP tasks such as Machine Translation and Semantic Recognition.

5. Results and Discussion

5.1. Results

In order to enable the monitor to display the color situation correctly, the angle 10° of CIE1964 was selected as the standard observer parameter, and the standard illuminant information was used as the D65 illuminant parameter. The tristimulus values calculated by the color integral equation were converted to the sRGB color gamut.

The spectral information provided by X-rite ColorChecker Classic was shown in Figure 6a. The result normalized by the standard diffuse reflection whiteboard was used to calculate the colors as shown in Figure 6b. The whiteboard with standard normalization is corrected by first derivative spectrophotometry, local similarity matching algorithm and centroid algorithm. The final calculated colors were shown in Figure 6d–f. As indicated, this type of correction algorithm was not applicable to spectrum color restoration. Through the hyperbolic tangent spectrum normalization and correction model with parameters proposed in this article, the results calculated by color integral are shown in Figure 6c.

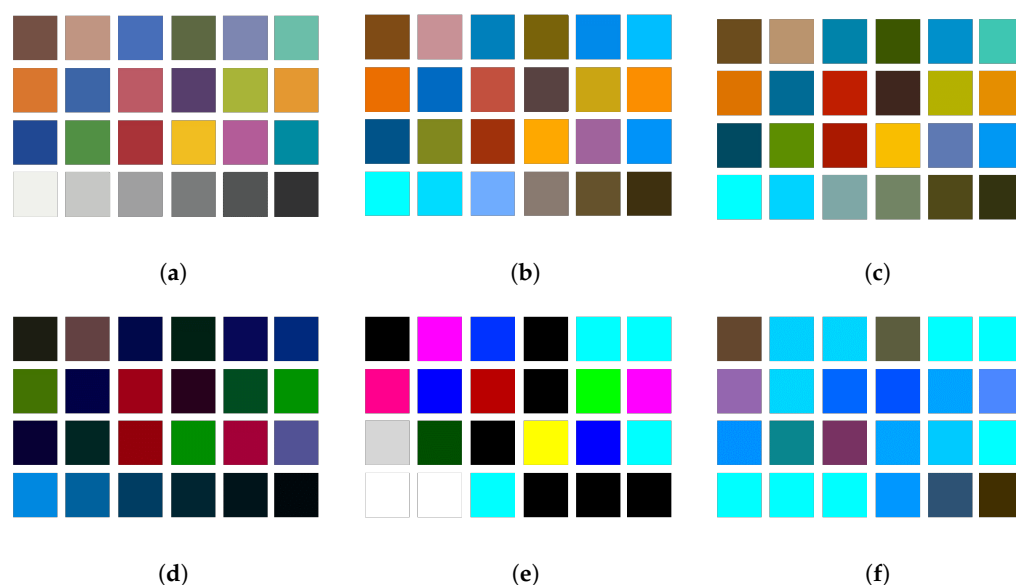


Figure 6. Color block diagram by color integral after normalization and correction. (a) X-rite ColorChecker Classic, (b) Standard diffuse reflection whiteboard results from normalization, (c) Hyperbolic tangent spectrum normalization and correction model with parameters, (d) First derivative spectrophotometry, (e) Local similarity matching algorithm, (f) Centroid algorithm.

According to the color calculation results, first derivative spectrophotometry, local similarity matching algorithm and centroid algorithm were not suitable for true-color reconstruction. The traditional normalization result based on diffuse reflection whiteboard and the spectral difference between the normalization optimization algorithm proposed in this article as well as the standard spectrum of X-rite ColorChecker Classic were shown in Figure 7, the calculation method is shown in Equation (10).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (s(\lambda_i) - \hat{s}(\lambda_i))^2} \quad (10)$$

Among them, the abscissa was the color number, and the ordinate was the spectral difference value. By comparison, it could be found that the spectrum after normalization and correction by the algorithm proposed in this article was closer to the standard spectrum.

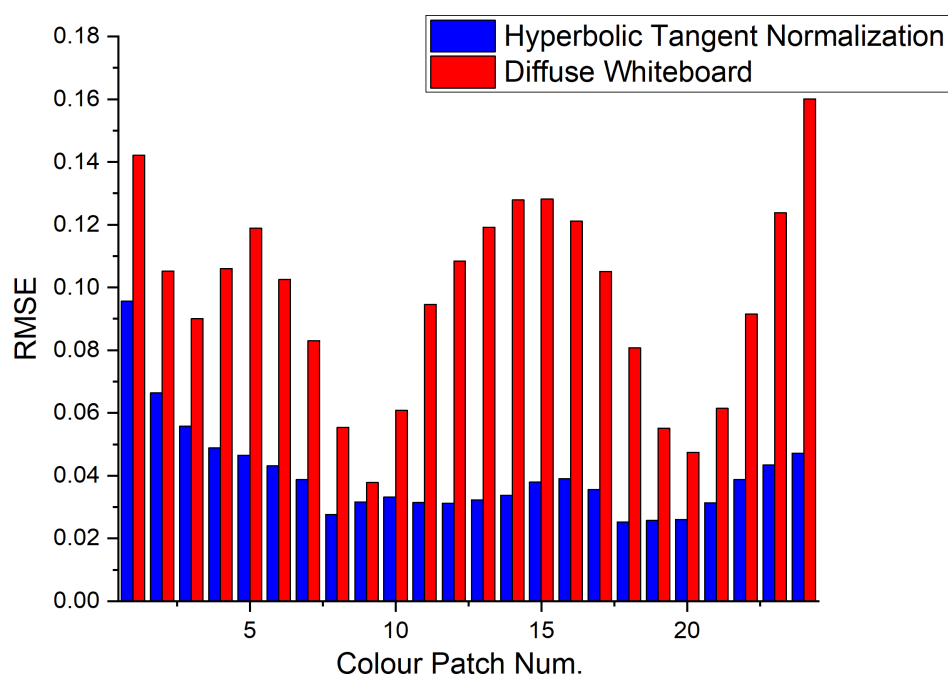


Figure 7. Spectral difference diagram between the traditional method and the algorithm proposed in this article with the standard spectrum.

According to the color calculation results in Figure 6, it can be found that there was obvious colour difference between the spectral information missing the 400–470 nm channel and the X-rite ColorChecker Classic. In this article, the processed ICVL multispectral data set was used as the priori data set. Since 12 multispectral natural images with resolution of were composed, there were a total of 21,715,200 spectral information in 400–700 nm band in the data set. Among them, the first 80% of the spectral information of each image was used as the priori data, and the last 20% of the spectrum data was used as the test data. We did a five-fold cross-validation [39] on the training set and test set. The two indicators for comparison are ΔE_{2000} , which mainly measures the color difference between the reconstructed color and the original color. The smaller the value, the more similar the two colors. The reconstructed part and the original part are calculated by RMSE, which measures the similarity between the reconstructed spectrum and the original spectrum. The smaller the value, the more similar the two spectra. The results are shown in Table 1.

Table 1. 5-fold cross-validation.

Num	ΔE_{2000}	RMSE %
		Only the Reconstruction Part Is Calculated
1	6.148465671	1.76555
2	6.061711912	1.591006
3	6.555953527	2.924124
4	6.503468054	2.899935
5	6.331956111	2.120579

Based on the calculated prior data parameters, the improved gradient boosting decision tree series forecasting algorithm was applied to reconstruct the missing channel between 400–470 nm, and the results were shown in Figure 8b. The bi-inverted Gaussian model is applied to reconstruct the missing part of spectrum, and the color calculation result was shown in Figure 8d. The color calculation results of the normalized spectrum based on the traditional standard diffuse reflection whiteboard was shown in Figure 8c. The color calculation results based on the Linear Interpolation was shown in Figure 8e.

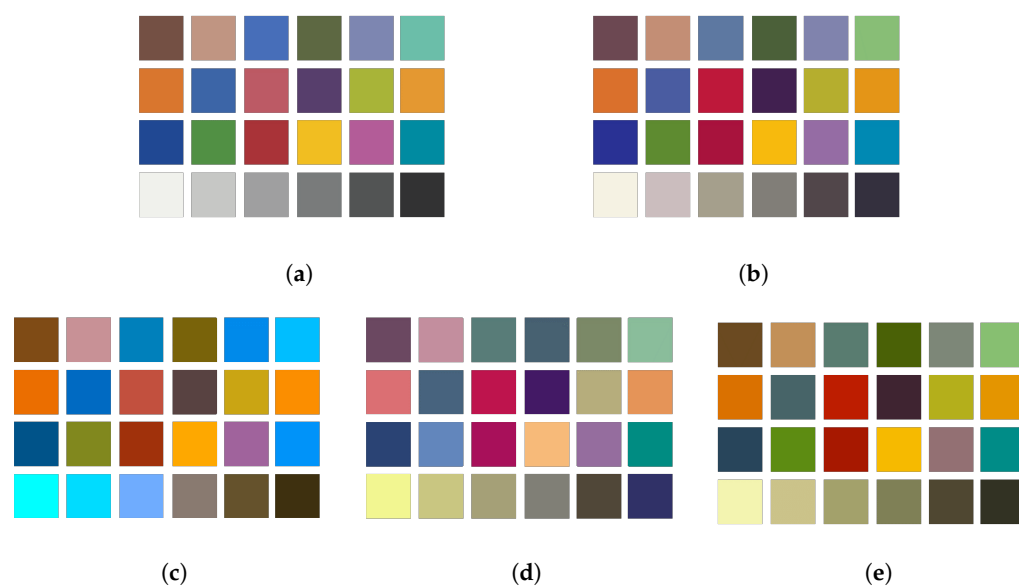


Figure 8. Color block diagram of spectral response function through color integral. (a) X-rite ColorChecker Classic, (b) Algorithm in this article, (c) Diffuse reflection whiteboard, (d) Bi-inverted Gaussian model, (e) Linear Interpolation.

Figure 9 showed the color difference between the reconstruction results based on the above three methods and the X-rite ColorChecker Classic. Among them, the abscissa was the color number, and the ordinate was the color difference value [40].

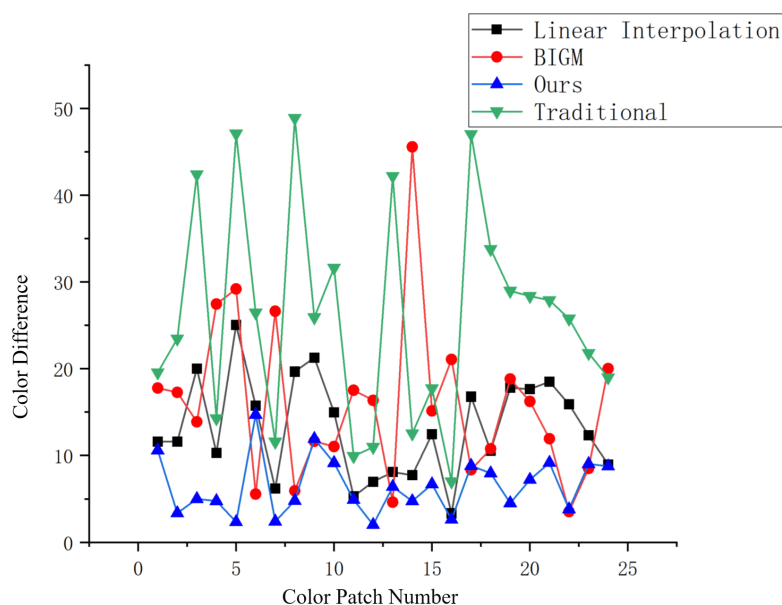


Figure 9. Color difference diagram between algorithm in this article, bi-inverted Gaussian model, Linear Interpolation, Ours, and Traditional method with standard spectrum.

According to the color difference value, there was smaller color difference between color information calculated by using the energy obtained by laser radar and X-rite ColorChecker Classic.

At the same time, a simulation experiment based on Munsell 1269 color chip was carried out to verify the adaptability of the algorithm in this paper. According to the spectral reflectance of each color published by Munsell, the color calculated by the color integral equation was converted into the CIE Lab Color Space. Figure 10a was obtained, where

value a was the abscissa (X -axis) and value b was the ordinate (Y -axis). After removing the 400–470 nm band information, the calculated result was shown in Figure 10b. Through improved spectrum reconstruction of gradient boosting decision tree series forecasting, the 400–470 nm band information was complemented, and the calculated results were shown in Figure 10c. The results obtained by complementing the spectrum by the bi-inverted Gaussian model and linear Interpolation were shown in Figure 10d,e respectively.

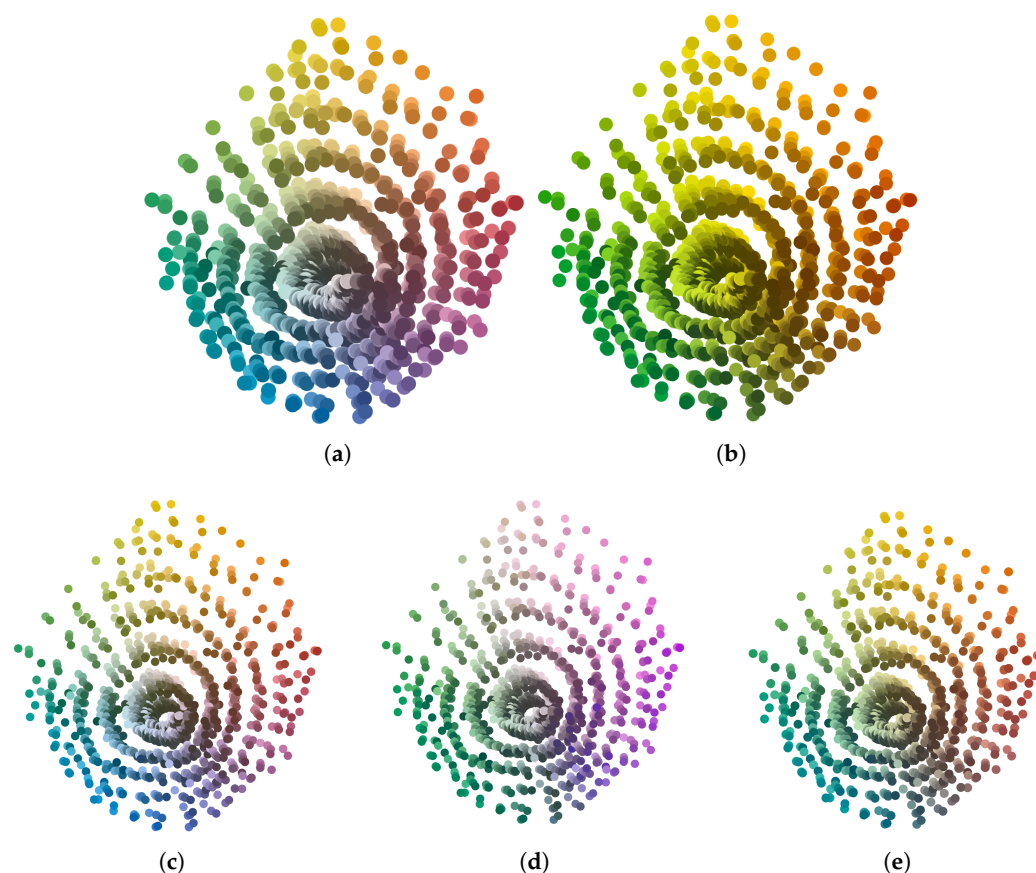


Figure 10. Munsell 1269 color chip simulation experiment. (a) Result of the complete spectrum, (b) Result of missing 400–470 nm band information, (c) Result of spectral recovery using the method proposed in this article, (d) Result based on bi-inverted Gaussian model, (e) Result based on linear interpolation algorithm.

The color difference relationship between the color calculated by the recovered spectrum based on the above 4 methods and the color calculated by the complete spectrum was calculated, as shown in Figure 11.

As can be seen from the line chart Figure 11, the average color difference between 24 colors and 1269 colors was calculated as shown in Table 2 for the color blocks missing the 400–470 nm band by the method proposed in this article.

As can be seen from Table 2, compared with B the bi-inverted Gaussian model and the linear interpolation algorithm, the color results calculated by the method proposed in this paper had the advantage of smaller color difference.

For both data results of Mean and Max, the performance of the improved algorithm will be significantly enhanced than that of the original method. In the analysis of Figure 11, there was relatively large color difference in individual colors. The color difference heat map Figure 12 involved the missing spectra in the 400–470 nm band Figure 12a, the restored spectra based on the algorithm Figure 12b in this paper, and the complete spectra. Since the missing 400–470 nm band belongs to the blue band, a serious color cast appeared in the blue color zone. The algorithm in this paper had a relatively large color difference between the cyan area and the central area, which caused large fluctuations in these parts of the

spectrum, which in turn led to the deviations in the prediction results. This could be used as the optimization direction for follow-up research.

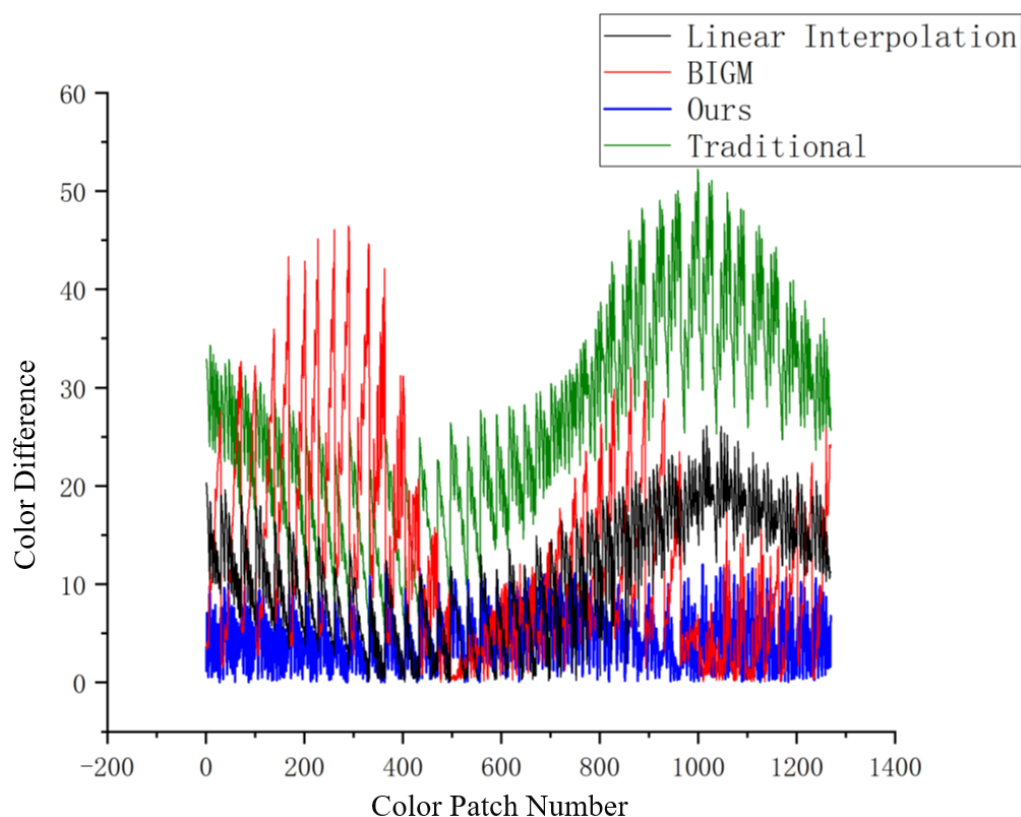


Figure 11. Color difference line chart of Munsell 1269 color chip.

Table 2. Color difference between X-rite ColorChecker Classic 24 color chip and Munsell 1269 color chip.

Num	Algorithm	Result	Color Difference of 1269 Munsell Color Chip	X-Rite Color Checker
1	Missing 400–700 nm	Mean	26.42	26.009
		Max	55.417	47.687
2	Algorithm in this paper	Mean	4.67	6.484
		Max	13.22472	10.514766
3	Bi-inverted Gaussian model	Mean	11.999	16.034
		Max	26.714542	22.414988
4	Linear Interpolation	Mean	11.052	13.28
		Max	29.2666469	21.1400417
5	Gradient Boosting Decision Tree	Mean	7.771	8.954
		Max	17.3334863	13.59832

In order to verify whether it can accurately restore the color information of geographic objects in the actual application, experiments on color card with 17 colors and real object are carried out, as shown in Figure 13.

By scanning the laser radar array of the color card with 17 colors, the spatial information of the color card with 17 colors and the energy information corresponding to the channel are obtained. The collected energy information is normalized and corrected through the hyperbolic tangent spectrum correction model with parameters. Then, spectral reconstruction is performed on the missing 400–470 nm bands based on the parameters calculated above. Finally, the spectrum integral of each reconstructed point is used to calculate the corresponding color information, as shown in Figure 14.

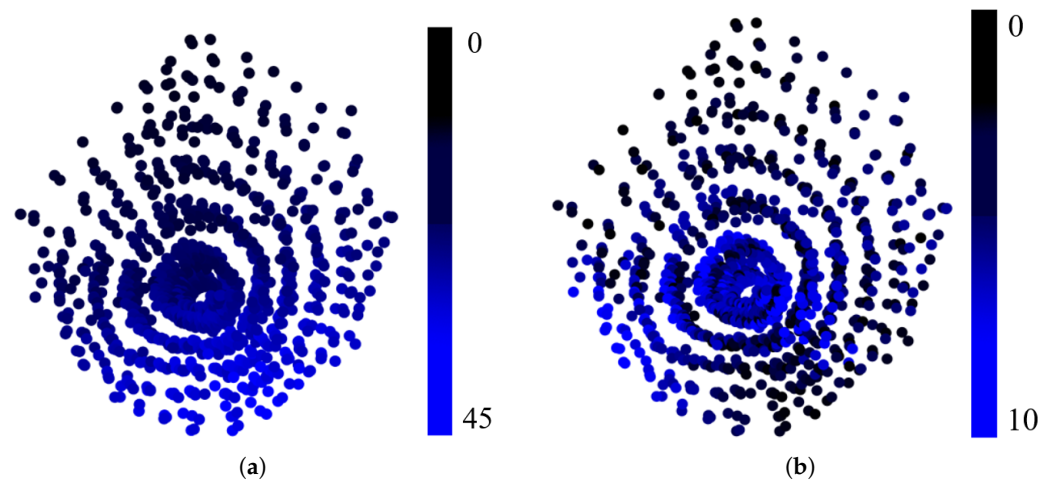


Figure 12. Normalized spectral response curve by whiteboard. (a) Result of missing 400–470 nm band information, (b) Result of spectral recovery using the method proposed in this article.



(a) Color card with 17 colors



(b) Real object

Figure 13. color card and real object.

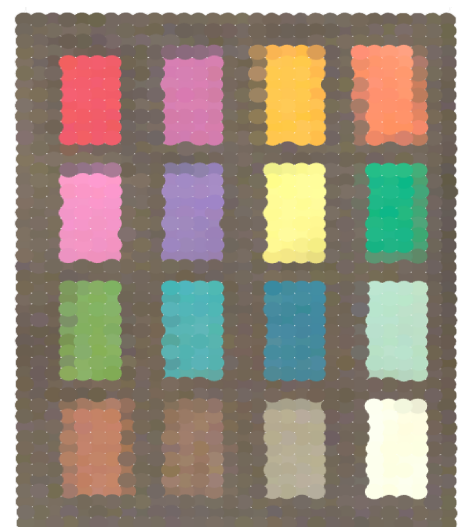


Figure 14. Result diagram of reconstruction on color card with 17 colors.

Based on the same approach, the result of real object data is shown in Figure 15. The experimental results show that the high spectrum correction algorithm and the missing

reconstruction algorithm proposed in this article can also restore the target color accurately in actual application.

Figure 15 is obtained by scanning the HSL system array. A total of 2800 points are obtained. In addition to the x.y.z coordinate data, there are 27 channels of data from 440–700 nm at 10 nm intervals.



Figure 15. Result diagram of reconstruction on real object.

Through the processing of the spectral reconstruction by hyperbolic tangent normalization and correction model with parameters, the reconstruction of the missing part by the improved gradient boosting decision tree sequence prediction algorithm, and the combination of the color value of the point calculated based on the color integral formula and the x.y.z coordinate data, the final result figure is obtained. In addition, the problem of mixed pixels also led to inaccurate colors. As shown in Figure 12, the red part of the Rubik's Cube appeared muddy. This issue was also worthy of in-depth study.

5.2. Discussion

The above experiment did not consider the negative impact of atmospheric effects on the high-precision acquisition of backscattered intensity. Various conditions such as atmospheric components will make the backscattered laser energy suffer a certain degree of attenuation. For some satellite LiDAR systems, atmospheric effects may cause varying degrees of interference to the reconstructed colors, which further causes the reconstruction accuracy to fail to reach the laboratory effect.

Similarly, good hardware performance of system configuration is a prerequisite for true color reconstruction. Continuous supercontinuum laser source device is needed, with spectral output range that can cover most of the visible range and the acquired signal that has a higher signal-to-noise ratio (SNR). In order to improve photoelectric transformation efficiency, photoelectric detectors also need to exhibit high quantum efficiency.

This research provides a reconstruction method that combines hyperspectral LiDAR echo energy and true-color. In the laboratory scenario, this method can more accurately reconstruct the true colors of color cards and objects in the absence of part of the visible light spectrum. Moreover, this method can also be applied to hyperspectral imaging to optimize the color reconstruction results.

6. Conclusions

This article proposes the hyperbolic tangent spectrum normalization and correction model and the improved spectrum reconstruction algorithm with gradient boosting de-

cision tree sequence prediction, with the aim of directly converting radar echo energy data into image color information. The point cloud information obtained by hyperspectral LiDAR and the energy information of corresponding channel are used as the experimental data set. The results show that the color integral equation based on the normalization and correction function and the missing part spectrum reconstruction provided in this article can accurately achieve the true color reconstruction. By comparing various spectrum correction algorithms based on 24-color and 1269-color cards, it is found that the traditional spectrum correction algorithm is not applicable to the true color reconstruction. Compared with the traditional normalization method with standard diffuse reflection whiteboard, the spectrum corrected by the algorithm proposed in this article is closer to that of the standard color card. The improved spectrum reconstruction algorithm with gradient boosting decision tree sequence prediction proposed in this article can accurately reconstruct the missing multispectral band information between 400–470 nm. In addition, experiments suggest that it can also achieve good effects in practical application of hyperspectral LiDAR. This article demonstrates the possibility of true color reconstruction through hyperspectral LiDAR echo energy data, which is of vital significance and value for multispectral imaging and ecological monitoring.

Author Contributions: Methodology, Visualization and Writing—original draft, T.W.; Resources, X.W. and S.S.; Data curation, B.C.; funding acquisition, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

GitHub: <https://github.com/wangtengfeng0809/True-color-Reconstruction-Based-on-Hyperspectral-LiDAR-Echo-Energy> (accessed on 1 May 2021).

References

1. Beger, R.; Gedrange, C.; Hecht, R.; Neubert, M. Data fusion of extremely high resolution aerial imagery and LiDAR data for automated railroad centre line reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, S40–S51. [\[CrossRef\]](#)
2. Debes, C.; Merentitis, A.; Heremans, R.; Hahn, J.; Frangiadakis, N.; van Kasteren, T.; Liao, W.; Bellens, R.; Pižurica, A.; Gautama, S.; et al. Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2405–2418. [\[CrossRef\]](#)
3. Khodadadzadeh, M.; Li, J.; Prasad, S.; Plaza, A. Fusion of hyperspectral and LiDAR remote sensing data using multiple feature learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2971–2983. [\[CrossRef\]](#)
4. Dong, J.; Zhuang, D.; Huang, Y.; Fu, J. Advances in multi-sensor data fusion: Algorithms and applications. *Sensors* **2009**, *9*, 7771–7784. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70–83. [\[CrossRef\]](#)
6. Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [\[CrossRef\]](#)
7. Sankey, T.; Donager, J.; McVay, J.; Sankey, J.B. UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA. *Remote Sens. Environ.* **2017**, *195*, 30–43. [\[CrossRef\]](#)
8. Gong, W.; Sun, J.; Shi, S.; Yang, J.; Du, L.; Zhu, B.; Song, S. Investigating the potential of using the spatial and spectral information of multispectral LiDAR for object classification. *Sensors* **2015**, *15*, 21989–22002. [\[CrossRef\]](#)
9. Liu, Q.; Liu, Y.; Pointer, M.R.; Huang, Z.; Wu, X.; Chen, Z.; Luo, M.R. Color discrimination metric based on the neutrality of lighting and hue transposition quantification. *Opt. Lett.* **2020**, *45*, 6062–6065. [\[CrossRef\]](#)
10. Liang, J.; Wan, X. Optimized method for spectral reflectance reconstruction from camera responses. *Opt. Express* **2017**, *25*, 28273–28287. [\[CrossRef\]](#)
11. Technical Committee International Commission on Illumination. *Practical Methods for the Measurement of Reflectance and Transmittance*; Commission Internationale de l’Eclairage: Vienna, Austria, 1998.

12. Liang, J.; Xiao, K.; Pointer, M.R.; Wan, X.; Li, C. Spectra estimation from raw camera responses based on adaptive local-weighted linear regression. *Opt. Express* **2019**, *27*, 5165–5180. [[CrossRef](#)] [[PubMed](#)]
13. Gong, W.; Song, S.; Zhu, B.; Li, F.; Cheng, X. Multi-wavelength canopy LiDAR for remote sensing of vegetation: Design and system performance. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 1–9.
14. Woodhouse, I.H.; Nichol, C.; Sinclair, P.; Jack, J.; Morsdorf, F.; Malthus, T.J.; Patenaude, G. A multispectral canopy LiDAR demonstrator project. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 839–843. [[CrossRef](#)]
15. Gaulton, R.; Danson, F.; Ramirez, F.; Gunawan, O. The potential of dual-wavelength laser scanning for estimating vegetation moisture content. *Remote Sens. Environ.* **2013**, *132*, 32–39. [[CrossRef](#)]
16. Song S.; Wang B.; Gong, W.; Chen Z.; Lin X.; Sun J.; Shi S. A new waveform decomposition method for multispectral LiDAR. *ISPRS J. Photogramm. Remote Sens.* **2019**, *3*, 40–49. [[CrossRef](#)]
17. Sun, J.; Shi, S.; Yang, J.; Chen, B.; Gong, W.; Du, L.; Mao, F.; Song, S. Estimating leaf chlorophyll status using hyperspectral lidar measurements by PROSPECT model inversion. *Remote Sens. Environ.* **2018**, *1*, 1–7. [[CrossRef](#)]
18. Zhao, X.; Shi, S.; Yang, J.; Gong, W.; Sun, J.; Chen, B.; Guo, K.; Chen, B. Active 3D imaging of vegetation based on multi-wavelength fluorescence LiDAR. *Sensors* **2020**, *3*, 935. [[CrossRef](#)] [[PubMed](#)]
19. Chen, B.; Shi, S.; Sun, J. Using HSI Color Space to Improve the Multispectral Lidar Classification Error Caused by Measurement Geometry. *IEEE Trans. Geosci. Remote Sens.* **2020**, *99*, 1–13. [[CrossRef](#)]
20. Wang, J.; Li, C. Development and prospect of hyperspectral imager and its application. *Chin. J. Space Sci.* **2021**, *41*, 22–33. (In Chinese)
21. Wang, B.; Song, S.; Gong, W.; Cao, X.; He, D.; Chen, Z.; Lin, X.; Li, F.; Sun, J. Color Restoration for Full-Waveform Multispectral LiDAR Data. *Remote Sens.* **2020**, *12*, 593–611. [[CrossRef](#)]
22. Chen, B.; Shi, S.; Gong, W.; Sun, J.; Chen, B.; Du, L.; Yang, J.; Guo, K.; Zhao, X. True-Color Three-Dimensional Imaging and Target Classification Based on Hyperspectral LiDAR. *Remote Sens.* **2019**, *11*, 1541. [[CrossRef](#)]
23. Zhou, J.; Zeng, Y.; Wang, X.; Wu, C.; Cai, Z.; Gao, B.Z.; Gu, D.; Shao, Y. The capture of antibodies by antibody-binding proteins for ABO blood typing using SPR imaging-based sensing technology. *Sens. Actuators B Chem.* **2019**, *304*, 1–15. [[CrossRef](#)]
24. Taulier, A.; Levillain, P.; Lemonnier, A. Determining methemoglobin in blood by zero-crossing-point first-derivative spectrophotometry. *Clin. Chem.* **2020**, *33*, 10–17. [[CrossRef](#)]
25. Owega, S.; Poitras, D. Local similarity matching algorithm for determining SPR angle in surface plasmon resonance sensors. *Sens. Actuators B Chem.* **2007**, *123*, 35–41. [[CrossRef](#)]
26. Wang, X.; Dong, L.; Zhan, S. Without-Baseline Centroid Algorithm for Surface Plasmon Resonance Spectra. *Chin. J. Sens. Actuators* **2012**, *25*, 365–369.
27. Zhan, S.; Wang, X.; Liu, Y. Fast centroid algorithm for determining the surface plasmon resonance angle using the fixed-boundary method. *Meas. Sci. Technol.* **2011**, *22*, 1–7. [[CrossRef](#)]
28. Adam, C.; Vivek, K. Denoising Hyperspectral Imagery and Recovering Junk Bands using Wavelets and Sparse Approximation. In Proceedings of the IEEE Geoscience and Remote Sensing Symposium, Denver, CO, USA, 31 July–4 August 2006; pp. 387–390.
29. Zhong, P.; Wang, R. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2260–2275. [[CrossRef](#)]
30. Yin, J.; Sun, J.; Jia, X. Sparse Analysis Based on Generalized Gaussian Model for Spectrum Recovery with Compressed Sensing Theory. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2752–2759. [[CrossRef](#)]
31. Liu, X.; Zhang, Y.; Teng, T.; Ding, Z. Estimation of Vegetation Water Content based on Bi-inverted Gaussian Fitting Spectral Feature Analysis Using Hyperspectral Data. *Remote Sens. Technol. Appl.* **2016**, *31*, 1075–1082.
32. Arad, B.; Ben-Shahar, O. Sparse Recovery of Hyperspectral Signal from Natural RGB Images. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 19–34.
33. Schonlau M. Boosted regression (boosting): An introductory tutorial and a Stata plugin. *Stata J.* **2005**, *5*, 330–354. [[CrossRef](#)]
34. Arumugam, P.; Kuppan, V. A GBDT-SOA approach for the system modelling of optimal energy management in grid-connected micro-grid system. *Int. J. Energy Res.* **2020**, *45*, 1–19.
35. Friedman J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2005**, *29*, 1189–1232.
36. Zhang, J.; Xu, D.; Hao, K.; Zhang, Y.; Chen, W.; Liu, J.; Gao, R.; Wu, C.; De Marinis, Y. FS-GBDT: Identification multicancer-risk module via a feature selection algorithm by integrating Fisher score and GBDT. *Brief. Bioinform.* **2020**, *1*, 1–13. [[CrossRef](#)] [[PubMed](#)]
37. Zhang, W.; Yu, J.; Zhao, A.; Zhou, X. Predictive model of cooling load for ice storage air-conditioning system by using GBDT. *Energy Rep.* **2021**, *7*, 1588–1597. [[CrossRef](#)]
38. Chen, Y.; Liu, Y. Which Risk Factors Matter More for Psychological Distress during the COVID-19 Pandemic? An Application Approach of Gradient Boosting Decision Trees. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1–17.
39. Amir, M.M.; Amirshahi, S.H. A step by step recovery of spectral data from colorimetric information. *J. Opt.* **2015**, *44*, 373–383. [[CrossRef](#)]
40. Cao, B.; Liao, N.; Cheng, H. Spectral reflectance reconstruction from RGB images based on weighting smaller color difference group. *Color Res. Appl.* **2017**, *42*, 327–332. [[CrossRef](#)]