A Point Cloud Optimization Method of Low Altitude Remote Sensing Image Based on Multi-channels

Lei Huang 1,3, Nan Yang 1,2,4, and Yu Zhang 2

1 Key laboratory of Urban Land Resource Monitoring and Simulation Ministry of Land and Resource of China, 8007 Hongli Road, Futian District Shenzhen China, 518034;
2 Harbin Institute of Technology, School of Transportation Science and Engineering, 73 Huanghe Road, Harbin, China, 150090;
3 Shenzhen Cadastral Surveying & Mapping Office, 69 Xinwen Road, Futian District Shenzhen China, 518034;
4 Shenzhen Research and Development Center of State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Keyuan Road, Shenzhen China, 518057.
Email: candyyanghot@126.com

Abstract. This paper presents a novel point cloud optimization method of low altitude remote sensing image based on multi-channels. The proposed method is designed to be especially effective for enhancing the accuracy of stereo matching. In order to fully exploit the information of multi-channel images, a weighted multi-channel intensity is employed instead of gray intensity. Then, an error equation is built to compute the optimal point according to the LSM method, space geometry relationship and collinear equation constraint. Compared with the traditional LSM method, the proposed method can achieve higher accuracy 3D point cloud data, since it can enrich the information conveys in the error equation and make the image matching robust. Comparison studies and experimental results prove the high accuracy of the proposed algorithm in low-altitude remote sensing image point cloud optimization.

1. Introduction

Using huge number of images to reconstruct a fine-detailed three-dimensional (3D) object or scene is more and more popular in various areas like urban planning [1], archaeology [2] agriculture [3], environmental protection [4], object tracking [5], and so forth. In the process of 3D reconstruction, one of the most significant procedures is image matching [6]. It is no doubt that the accuracy of image matching can directly ordain the quality of the reconstructed 3D model [7]. Accordingly, numerus researchers have spent nearly half a century period to improve the accuracy of the image matching, like Least Square Matching (LSM) in 1980s [8], Multi-photo Geometrically Constrained (MPGC) in 1990s [9], Least Squares B-Spline Snakes (LSB-Snakes) in 2000s [10], Patch-based Multi-View Stereo (2010s) [11] and so forth. However, although a great deal of image matching refinement algorithms had greatly enhanced the precision of image matching, noise; similar problems are still encountered in most matching refinement approaches [12]. The main problems in stereo matching are encountered with the following:

- Object discontinuities;
- Occlusions;
- Little or no texture.
With the development of technologies about aerial photography as well as information storage and communication, color or multi-channel images can be achieved more and more easily, and the amount of information which people can get from the modern images has been improved. Since grey images is always used in the most matching refinement algorithms, multi-channel images must be transmuted into grey images, hence much information in the multi-channel images has been lost. The proposed method introduces multi-channel intensity in the images to the matching refinement process. By using multi-channels intensity of the pixels into the LSM error equations instead of grey intensity to enlarge the information of the images and enhance the stability of image matching. Low-altitude remote sensing image data at Northwestern University Campus and Hainan district are used to verify the feasibility of this method. The experiments illustrate that the matching optimized method can effectively raise the rate of successful matching and improve the matching precision of 3D points.

The remainder of the paper is structured as follows: the proposed method is introduced in Section 2; in Section 3 the experiments are developed to verify the feasibility of the method in terms of reliability and matching accuracy; lastly, the conclusions are stated in Section 4.

2. Method
In this section, the point cloud optimization method based on multi-channels is developed. The proposed method aims at raising the stability of image matching and improving the matching accuracy. The workflow of this algorithm is shown in Fig.1. The proposed method employs three steps to optimize the object points in point cloud data: 1) Construction of error equation; 2) calculation of pixel intensity; 3) Outliers filter. By introducing multi-channels intensity of the pixels in the matching refinement process, the robust and accuracy of image matching is improved.
2.1. Construction of Error Equation

The main idea of matching refinement algorithm is to find a pair of image windows in the reference image and search image. When the value of Normalization Cross Correlation (NCC) between the two image windows reaches the maximum, the two central pixels in the corresponding image windows could be the corresponding pixels.

According to the LSM method, the matching equation is:

\[
v = dh_0 + g_i(x_i, y_i) \cdot dh_i + h_i \left( \frac{\partial g_i}{\partial x_i} dx_i + \frac{\partial g_i}{\partial y_i} dy_i \right) - \left( g_i(x_0, y_0) - h_0 - g_i(x_i, y_i) \right)
\] (1)

In the error equation above, \(v\) is the projection error, \(h_0\) and \(h_i\) are the radiation distortion coefficients between the reference image and search image \(i\). \(dh_0\) and \(dh_i\) are corrections of parameter \(h_0\) and \(h_i\). \(g(x_0, y_0)\) and \(g(x, y)\) are intensity value of corresponding pixels \((x_0, y_0)\) and \((x, y)\) which are located in the image window of the reference image and search image \(i\) respectively. \((\partial g_i/\partial x_i, \partial g_i/\partial y_i)\) is the derivative value of pixel intensity in the \(x\) and \(y\) direction. \((dx_i, dy_i)\) is the correction values of the image points \((x_i, y_i)\). Since the purpose of the present method is to optimize the point cloud, according to the collinear equation, \((x_0, y_0)\) and \((x, y)\) can be represent as:

\[
x_{(0/i)} = -f_{(0/i)} \begin{vmatrix} a_{(0/i)} & a_{2(0/i)} & a_{3(0/i)} \\ b_{(0/i)} & b_{2(0/i)} & b_{3(0/i)} \\ c_{(0/i)} & c_{2(0/i)} & c_{3(0/i)} \end{vmatrix}
\]

\[
y_{(0/i)} = -f_{(0/i)} \begin{vmatrix} a_{2(0/i)} & a_{3(0/i)} & a_{3(0/i)} \\ b_{2(0/i)} & b_{3(0/i)} & b_{3(0/i)} \\ c_{2(0/i)} & c_{3(0/i)} & c_{3(0/i)} \end{vmatrix}
\] (2)

In the collinear equation above, \((X, Y, Z)\) is the coordinate of object point in the object space coordinate system, \((X_{00}, Y_{00}, Z_{00})\) and \((X_{i0}, Y_{i0}, Z_{i0})\) are the coordinates of the two images’ projection center in the object space coordinate system, \(R_{00}\) and \(R_i\) are rotation matrix of reference image and search image \(i\).

Since a pair of corresponding pixels can build only one error equation and it is impossible to seek solution by only one error equation, a pair of image matching windows which contain several pixels surrounding the pair of corresponding pixels is essential, moreover, the pixels which are located at the same position of the reference image and search image windows respectively are theoretical corresponding pixels. More than one error equation therefore can be built, and if the number of error equations is larger than the number of unknown quantities, least square adjustment can be employed to get more accuracy result.

2.2. Calculation of Pixel Intensity

Generally, color images or multi-channel images should be changed into grey images in the matching progress. The intensity value of the pixel can be calculated as:

\[
g(x, y) = \frac{1}{n} \sum_{k=1}^{n} I_k(x, y)
\] (3)

\(I_k(x, y)\) is the intensity value of pixel \((x, y)\) in channel \(k\).

The proposed method introduces a weighted algorithm instead of arithmetic mean to calculate the intensity value which can be calculated as:
Because the interval of the adjacent pixels is extremely small, the difference between neighboring pixels is employed to calculate $\frac{\partial g(x_i, y_i)}{\partial x_i}$ and $\frac{\partial g(x_i, y_i)}{\partial y_i}$ in the error equation (1) instead of differential. The derivative value of pixel intensity in the $x$ and $y$ direction can be approximately expressed as:

$$\frac{\partial g_i}{\partial x_i} = \frac{1}{2} \left( g_i(x_i + 1, y_i) - g_i(x_i - 1, y_i) \right)$$

$$\frac{\partial g_i}{\partial y_i} = \frac{1}{2} \left( g_i(x_i, y_i + 1) - g_i(x_i, y_i - 1) \right)$$

Simultaneously introducing formula (2), (4), (5) to error equation (1), using $\mu \times \mu$ pixels in image window and employing least square adjustment to calculate the premium matching point and optimize the coordination of 3D point.

### 3. Experiments and Results

In this section, two sets of point cloud data are used in the experiment. The data is generated from two pair of Unmanned Aerial Vehicle (UAV) images captured in Northwestern University Campus (Xi’an, Shanxi province) and Hainan province respectively. Detailed information of the images and initial point cloud are illustrated in Table.1, Table.2, Figure.2 and Figure.3.

**Table 1. Parameters of the Photography in Northwest University (46337 points)**

| Camera Name | CCD Size (mm) | Image Size (pixel) | Pixel size (μm) | Focal Length (mm) | Flying Height (m) | Ground Resolution (m) |
|-------------|---------------|-------------------|----------------|------------------|------------------|-----------------------|
| Canon EOS 400D | 22.16×14.77 | 3888×2592 | 5.7 | 24 | 600 | 0.118 |

**Table 2. Parameters of the photography in Hainan (70439 points)**

| Camera Name | CCD Size (mm) | Image Size (pixel) | Pixel size (μm) | Focal Length (mm) | Flying Height (m) | Ground Resolution (m) |
|-------------|---------------|-------------------|----------------|------------------|------------------|-----------------------|
| Canon EOS 5D | 36×24 | 5616×3744 | 6.4 | 24 | 650 | 0.174 |
3.1. Parameters Setting
In the proposed method, the interior and exterior orientation parameters of the images are known, and
the initial matching is the reprojection from initial 3D point cloud to reference and searching images.
To solve error equations, an iteration is used that is set to stop when the parameter corrections $dh_{0i}$ and
$dh_{1i}$ are less than $10^{-5}$ other parameter corrections are all less than $10^{-3}$ or the iteration is more than 300.
The matching results with image window size of $7 \times 7$, $11 \times 11$, $13 \times 13$, $15 \times 15$, $17 \times 17$, $19 \times 19$ and $21 \times 21$
are compared, respectively.

In the process of individual point matching, the Normalized Correlation Coefficient can reflect the
degree of approximation between image windows. Theoretically, the NCC is closer to 1, the
correlation between image windows is higher, and the center pixels of the two image windows are
more probable to be a pair of corresponding points. Therefore, when the NCC value between the
image matching windows is larger than a fixed threshold, the center pixels can be considered as valid
matching, and their corresponding object point on the ground can be added in the point cloud. In the
experiments, an NCC value 0.6 is set as threshold.

Moreover, to improve the accuracy and reduce the number of outliers in point cloud, the proposed
method enforces surface prior constraint to remove erroneous points. A common density constraint24
is employed to filter the outliers in point cloud. At last, the elevation difference between proposed
method and patch-based LSM method is compared to estimate the accuracy of the point cloud after
outliers filter step.

3.2. Result and Analysis
For stereo matching with low altitude images by photography, the error of altitude is the largest, hence
the altitude difference is always considered as the criteria to evaluate the accuracy. The experiments
employed altitude difference between the difference between the Z-coordinates calculated after the

Figure 2. Initial Point Cloud of Northwestern University Campus

Figure 3. Initial Point Cloud of Hainan Area
end of the iteration and the Z-coordinates calculated by the forward intersection. The comparison performances between patch-based LSM and the proposed method are illustrated in Fig.4 and Fig.5.

![Figure 4](image1.png)

**Figure 4.** Percentage of Successful Matching: (a) Northwestern University; (b) Hainan

![Figure 5](image2.png)

**Figure 5.** Altitude Difference Comparison: (a) Northwestern University; (b) Hainan

Fig.4 compared the percentage of valid points after filter step. The line chart indicate that the percentage of successful matching increased along with the expansion of the matching window size. It is because that more pixels involved in the matching process can decrease the weight of the noise and enhance the stability of image matching. It also explained the reason for which the proposed method had higher successful rate. Moreover, when the matching window size reached 19×19 or more, the percentage of successful matching had almost the same amount in spite of different methods since the error equations contain sufficient information to guarantee the stability of image matching.

In the line charts of Fig.5, X-axis represent matching window size and Y-axis is the altitude difference. By comparing the altitude difference between LSM and the proposed method in each chart, the results can be demonstrated that the error of proposed method is lower than LSM method, so that the proposed method which introduce multi-channel information to the image matching progress can improve the accuracy of stereo matching. From the two charts, the matching accuracy of Northwestern University images is superior to Hainan images due to the higher ground resolution and simpler image texture. Moreover, the Northwestern University images produced more effect in accuracy improvement, it probably because this image sets are more colorful than the Hainan images.

4. Conclusion
In this paper, we presented a novel method for improving the accuracy of point cloud reconstructed by low altitude remote sensing images. The propose method employs a weighted multi-channel intensity instead of arithmetic mean to build error equations. The purpose of this method is utilize informative pixel intensity to raise the stability and accuracy of stereo matching. Meanwhile, the matching results and the accuracy of the algorithm are validated experimentally.

The comparison of successful matching and altitude difference has demonstrated that the proposed approach is superior to the patch-based LSM method. It is worth noting that the effectiveness of the proposed method is depended on the variety of image shade. It is better used for matching images with
manifold channels and colors. A possible future work is matching images which has near-infrared channels as vegetation has higher reflection rate in these channels, which can distinguish vegetation from manifold artificial structures and objects to decrease the error matching from occlusion caused by plants.

5. Acknowledgments
The project was supported by the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Land and Resources (KF-2018-03-026).

6. References
[1] Kelly, T., Femiani, J., Wonka, P., & Mitra, N. J. (2017). Bigsur: large-scale structured urban reconstruction. ACM Transactions on Graphics, 36(6), 1-16.
[2] Hausmann, J., Zielhofer, C., Werther, L., Berg-Hobohm, S., Dietrich, P., & Heymann, R., et al. (2017). Direct push sensing in wetland (geo) archaeology: high-resolution reconstruction of buried canal structures (fossa carolina, germany). Quaternary International, S1040618216308291.
[3] Jiang, S., & Jiang, W. (2017). Efficient structure from motion for oblique uav images based on maximal spanning tree expansions. Isprs Journal of Photogrammetry & Remote Sensing.
[4] Nico, M., & Tobias, B. (2017). Structure-from-motion using historical aerial images to analyse changes in glacier surface elevation. Remote Sensing. 9(10), 1021-.
[5] Zheng, C., & Wei, Z. (2017). Object tracking by transitive learning using perspective transformation with asymptotic stability. Journal of Applied Remote Sensing, 11(4), 1.
[6] Knapitsch, A., Park, J., Zhou, Q. Y., & Koltun, V. (2017). Tanks and temples: benchmarking large-scale scene reconstruction. ACM Transactions on Graphics, 36(4), 1-13.
[7] Yong, Z., Xiuxiao, Y., Yi, F., & Shiyu, C. (2017). Uav low altitude photogrammetry for power line inspection. ISPRS International Journal of Geo-Information, 6(1), 14-.
[8] Ackermann, F. (1984). Digital image correlation: performance and potential application in photogrammetry. Photogrammetric Record, 11.
[9] E.P. Baltsavias. 1991. Multiphoto Geometrically Constrained Matching. Diss. Techn. Wiss. ETH Zürich, Nr. 9561, 1991. Ref.: A. Gruen; Korref. : BP Wrobel.
[10] L. Zhang. (2005). Automatic Digital Surface Model (Dsm) Generation from Linear Array Images. Mitteilungen- Institut fur Geodasie und Photogrammetrie an der Eidgenossischen Technischen Hochschule Zurich.
[11] Furukawa, Y. (2010). Accurate, dense, and robust multiview stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32.
[12] Ozyesil, O., Voroninski, V., Basri, R., & Singer, A. (2017). A survey of structure from motion. Acta Numerica, 26.