

3D-Rebuild based SFM Algorithm Improvement

Junyuan Sun
Jiangsu University
Jiangsu University, Zhenjiang, Jiangsu, China
Nankai District, Tianjin, China

ABSTRACT

The reconstruction of 3D scenes requires graphic processing algorithms. Different results are given and compared while different processing algorithms are injected during each stage of corner points processing, and the adaptive corner point algorithms are selected. In this work, this paper provided a novel reconstruction algorithm using different comparison methods. This method displays excellent matching accuracy, its dot-matrix is quickly converged, and the result demonstrate high fidelity. The algorithm can avoid the scene to be simple or fake, thus can be applied to reconstruct 3D objects.

General Terms

Digital Graphic Processing

Keywords

Digital Graphic Processing, 3D Reconstruction, KLT algorithm.

1. INTRODUCTION

With the development of computer hardware and the improvements on algorithms, 3D reconstruction method could be gradually used on many areas, including amplified reality, large scene construction and other technologies. In the real world, 3D reconstruction is the core technology in computer graphics and virtual reality, as processing information from two-dimensional graphics is low-cost, the information is easy to acquire, and the algorithm is independent to hardware.

SFM method is the most widely applied methodology, and the corner point algorithm is the most important step in SFM

reconstruction method. Nowadays, the major corner point methods include SIFT (Scale-invariant feature transform) algorithm, ORB (Oriented FAST and Rotated BRIEF) algorithm, SURF (Speeded Up Robust Features) algorithm, and KLT (Kanade-Lucas-Tomasi Tracking) algorithm. In this word, experiment compared these results using the above listed algorithms, selected the one with the most accurate result, and improve the selected algorithm.

Herein, chose the most robust SIFT algorithm to extract the matching points of each graphics, and deployed KLT feature following algorithm to renew the matching data of the corner points. Control groups were set up to compare and analysis the data, which prove that using the method in this article can offer a more realistic reconstruction result.

2. SFM Algorithm

SFM (Structure from Motion) algorithm is an algorithm using pictures shot from different angles and various distances to determine the structural relationships between disparate reference frames and individual graphics. When building a 3D model with SFM algorithm from the pictures, users can use handheld devices such as smartphones and card cameras rather than professional RGB cameras to acquire the requested pictures, and the algorithm is adaptive in many environments such as the construction of large scenes and the fusion of small scenarios.

The workflow of scene reconstruction deploying SFM algorithm is shown in figure 1. In this work, each step in the workflow is introduced and focus on comparing the reconstruction results of different algorithms.

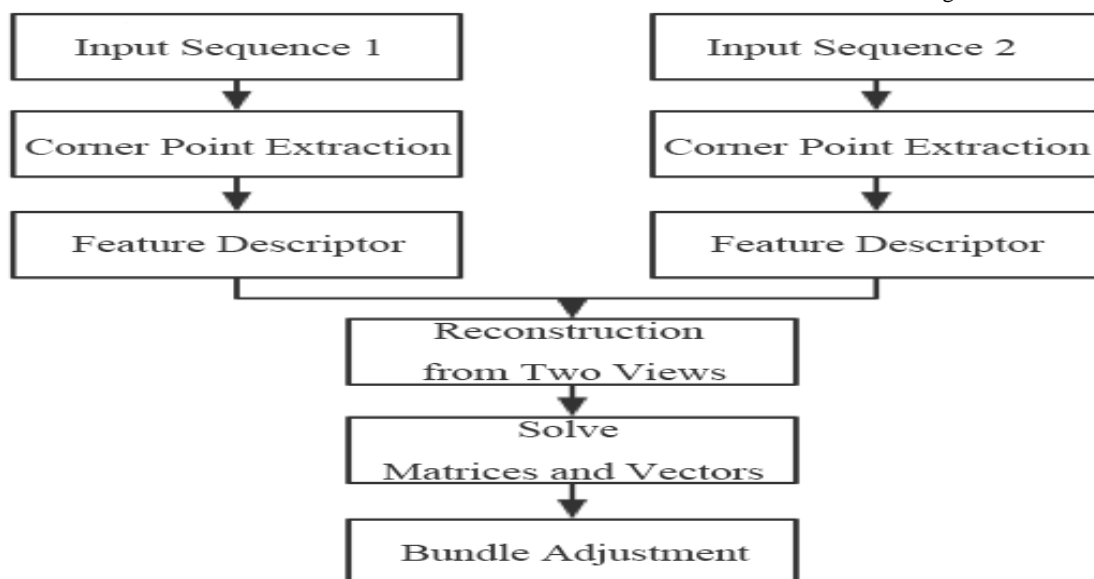


Fig 1: SFM workflow

3. CORNER POINT EXTRACTION AND FEATURE DESCRIPTOR

There major corner point extraction and matching algorithm includes ORB, SIFT, and SURF. By analyzed the number of corner points (figure 2A) detected and the average time of detection (figure 2B), which suggested that although SURF algorithm is an upgrade to SIFT algorithm, the upgrade focus

only on decreasing the time complexity, which means the corner points could be faster detected, as shown in figure 2B, but displayed worse robustness than SIFT when the illumination is changed or the graphic is distorted. As the most important question is how to ensure the accuracy of corner point extraction and matching, SIFT seemed to be better than SURF algorithm

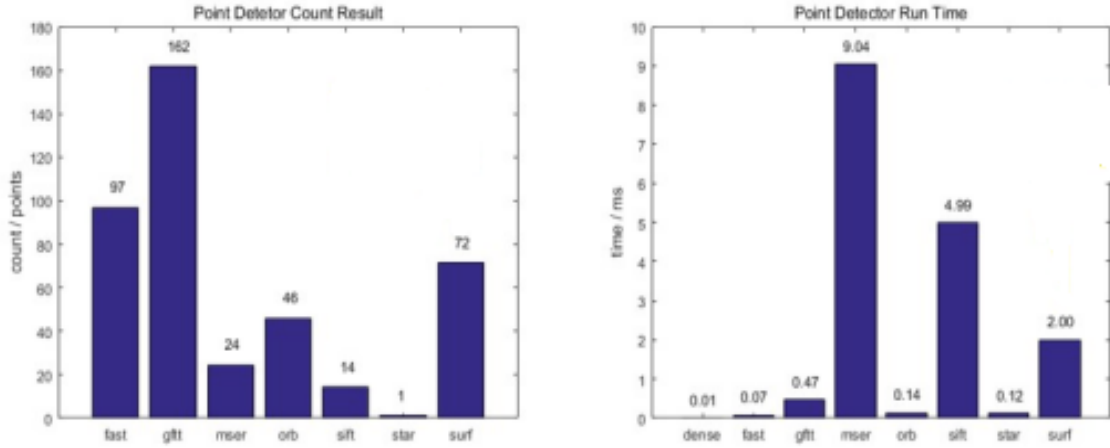


Fig2.: Comparison Between Different Algorithms

3.1 SIFT Algorithm

SIFT was initially reported by David Lowe in 1999 as an algorithm digging the local features of graphics by selecting the corner points from the built scale space, to acquire the spatial and other scale information from the built scale of the graphic.

The realization of this algorithm can be divided to three steps listed below

3.1.1 Constructing Scale Space

According to the specific graphical input, the algorithm is initially deployed to determine the parameters of its scale space.

The scale space $L(x, y, \sigma)$ is the convolution between the Gaussian function of the scale of change $G(x, y, \sigma)$ and the original figure $L(x, y)$

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Herein, suggest the dimension of Gaussian function is m and n, the pixel position is (x,y) and the scale space factor is σ , 12.5% convolution can be derived from the following formula:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m/2)^2 + (y-n/2)^2}{2\sigma^2}}$$

3.1.2 Constructing Gaussian Pyramid

Proposed that the size of the figure is m*n, according to the pyramid model, the minimum of the top dimension should initially be calculated its logarithm is defined as t, as listed below:

$$n = \log_2 \{ \min(M, N) \} - t, t \in [0, \log_2 \{ \min(M, N) \}]$$

The operation of extreme value detection requires the calculation of the scale space coordinate σ , the number of groups O, the inner layer index s, and the number of layers S, the construction formula is as follows:

$$D(x, y, \sigma) = (G(x, y, \sigma(s+1)) - G(x, y, \sigma(s))) * I(x, y) \\ = L(x, y, \sigma(s+1)) - L(x, y, \sigma(s))$$

(Initial scale = 1.6, S = 3)

3.1.3 Eliminating Edge Responsive Points

The procedure of eliminating edge responsive points from Hessian matrix is listed below:

α and β represent the direction gradient of x and y respectively. In this case, α must be larger than β , and set a parameter γ to satisfy $\alpha = \gamma \beta$.

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta,$$

$$\text{Det}(H) = D_{xx} D_{yy} - (D_{xy})^2 = \alpha \beta$$

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$$

The goal is to let $\frac{(\gamma + 1)^2}{\gamma}$ as small as possible, so we only have to test whether each corner point satisfies

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(\gamma + 1)^2}{\gamma}$$

3.2 KLT Algorithm

Gibson proposed the KLT method in 1950 to determine the relationship between two frames by calculating the correlating

information of the frames over time in the motion map set, and finally sort out the motion of the object during the action.

The algorithm is mainly embodied in three different forms depending on the scene (based on region and feature, frequency domain based, and gradient based).

3.2.1 Optical Flow Algorithm

While using the optical flow algorithm, this time-consuming extraction process of the so-called "good feature" can be omitted by calling the dense optical flow of the image using the global optical flow method.

Firstly, using the local optical flow method can extract the feature points, then calculate the optical flow, and reasonably use the sparse optical flow of the feature points to reflect the dense optical flow of the entire scene. This method has high accuracy and computational efficiency compared to other algorithms.

The feature-based optical flow method usually extracts the required feature points and calculates the optical flow in the first picture (the first frame). Based on the above requirements, the extracted feature points must be guaranteed to provide representative information, and with certain accuracy.

3.2.2 Feature Pyramid Network

Recursively solve and calculate the data of the original image from the highest layer (the lowest resolution picture) of the established pyramid. Here we set the L layer loss function to:

$$\epsilon^L(d^L) = \epsilon^L(d_x^L, d_y^L) = \sum_{x=u_x^L-\omega_x}^{u_x^L+\omega_x} \sum_{y=u_y^L-\omega_y}^{u_y^L+\omega_y} (I^L(x, y) - J^L(x + g_x^L + d_x^L, y + g_y^L + d_y^L))^2.$$

Iterative Lucas-Kanade algorithm:

The velocity vector is v at the k th iteration; the $k+1$ th solution is also using the standard Lucas-Kanade algorithm, paying attention to the need to update the neighborhood of the image at point $[px, py]$ in the process:

$$\forall (x, y) \in [p_x - \omega_x, p_x + \omega_x] \times [p_y - \omega_y, p_y + \omega_y], B_x(x, y) = B(x + v_x^{k-1}, y + v_y^{k-1})$$

The solution of $\mathbf{g}_x^L \mathbf{g}_y^L$ can be obtained by the formula:

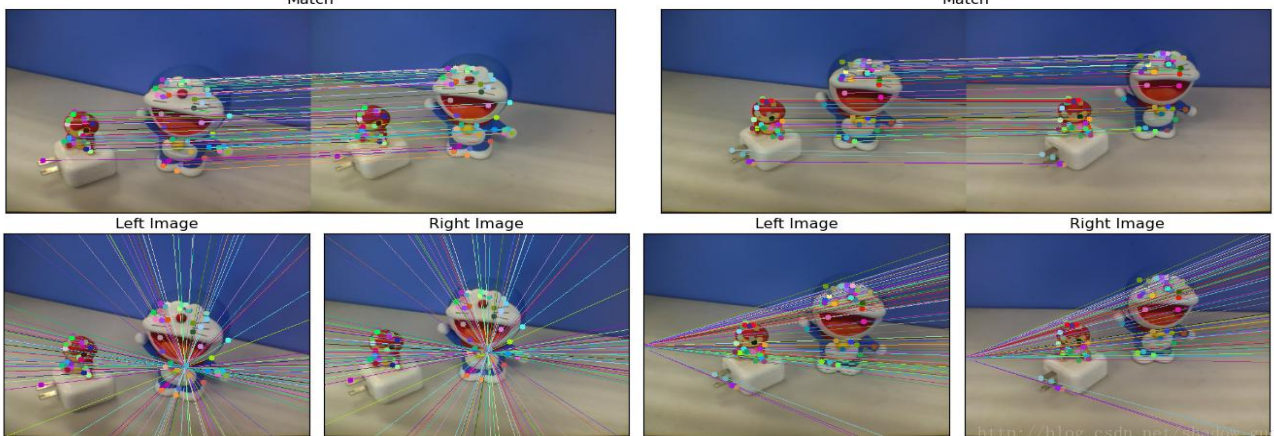


Fig3: Experiment Result 1.1

$$\mathbf{g}^L = 2(\mathbf{g}^{L+1} + \mathbf{d}^{L+1})$$

$$\bar{v} = d^L = \bar{v}^{-K} = \sum_{k=1}^K \bar{\eta}^{-k}$$

Thus we would finally solve v :

4. Bundle Adjustment

The beam adjustment method (BA operation) optimizes the motion matrix of multiple cameras in the established space coordinate system. They are used to handle data loss situations, provide true maximum likelihood estimation, and reconstruct 3D models in motion.

Cost function:

$$\arg \min_{P_i, M_i} \sum_{i=1}^k \sum_{j=1}^n v_{ij} d(Q(P_j, M_i), m_{ij})^2$$

(kk space points; nn number of imaging planes; feature point coordinates corresponding to the i -th point of m_{ij} on the j th imaging plane; whether v_{ij} point i has a projection on plane j ; outer parameter vector of P_i imaging plane; M_i three-dimensional point coordinates Vector; $Q(a_j, b_i)$ reprojection function; $d(x, y)$ distance metric function)

5. Result Comparison

5.1 Single SIFT Algorithm

- 1) Feature point extraction is relatively stable, the number is relatively large, and its robustness during motion is high.
- 2) When the 3D model is built, SIFT processes the image faster.
- 3) The amount of feature information that can be given is sufficient, and it can be guaranteed to be fast and accurate in the face of very large data.
- 4) Poor time continuity.
- 5) Unstable brightness characteristics.

The result is depicted in figure 3.

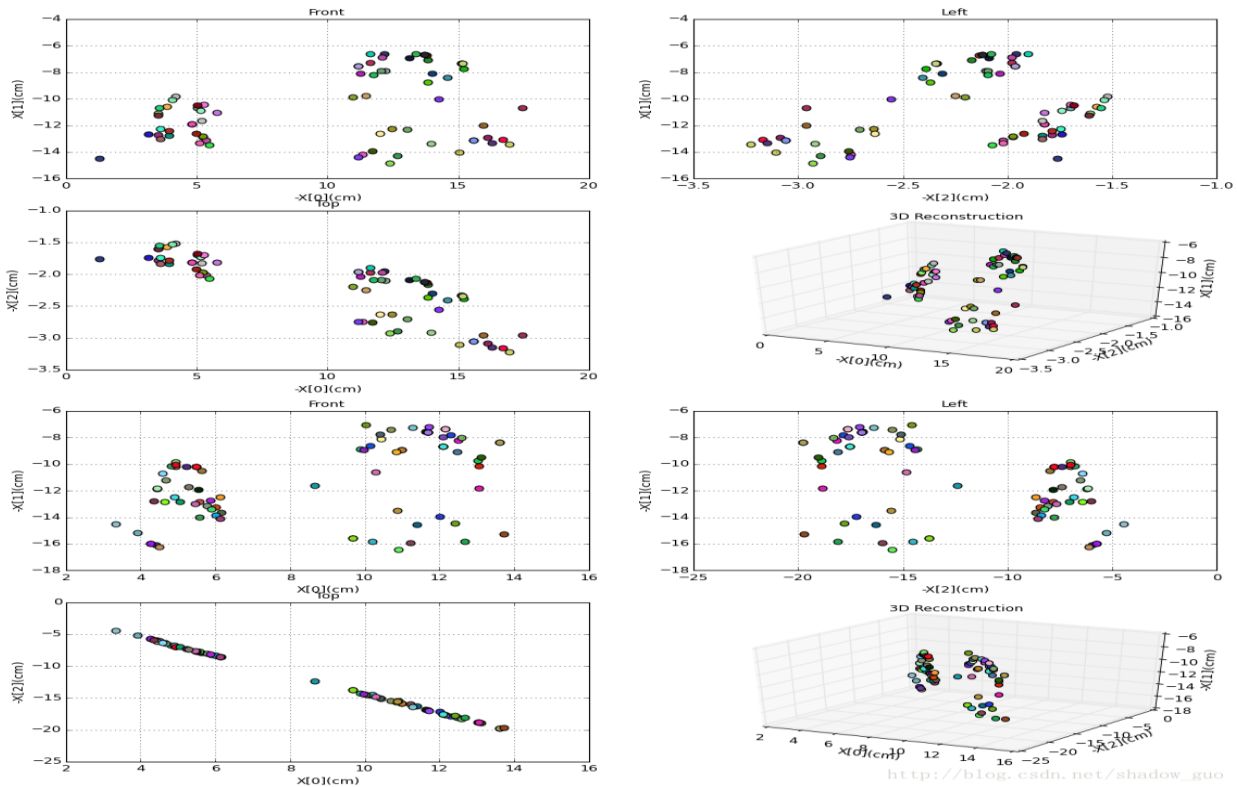


Fig4: Experiment Result 1.2

The illustration shows the result of the reconstruction. The actual data of the pendant is 10 cm high and 7 cm wide. The reprojection error is 1.058632472 and 8.451836329 pixels.

5.2 SIFT+KLT

- 1) Feature points have scale invariance.
- 2) Feature points are highly representative
- 3) The effective feature points extracted by the optical flow equation can be tracked by optical flow and have locality

and affine invariance after reconstruction.

Constant brightness, continuous time axis and small displacement of motion

- 4) Spatially consistent.

Improved results. There is no big difference in the feature point matching algorithm. The motion of the four pictures selected was small.

The result is depicted in figure 4.

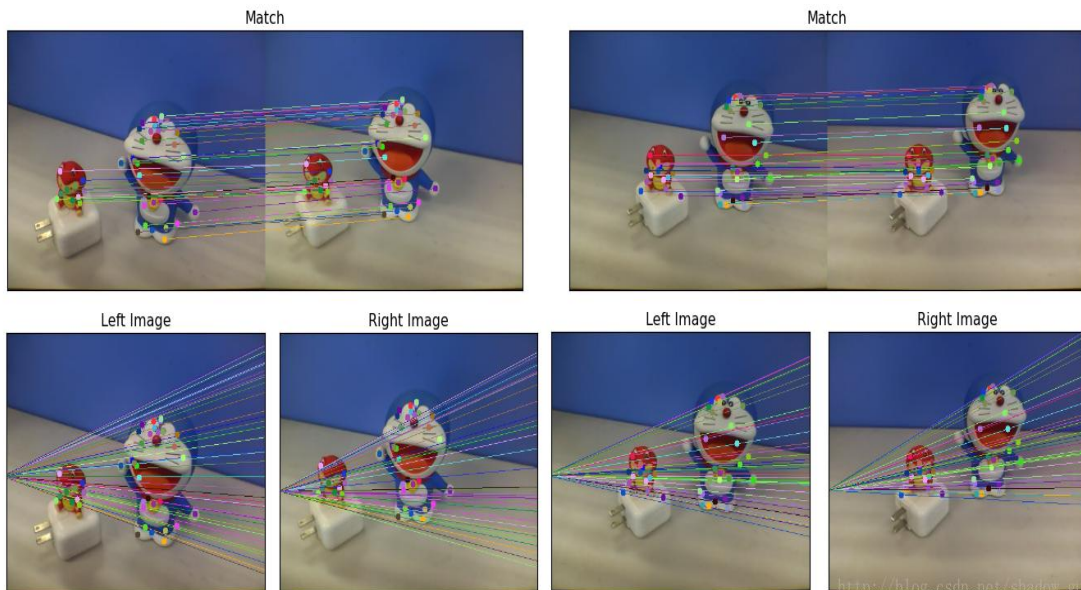


Fig5: Experiment Result 2.1

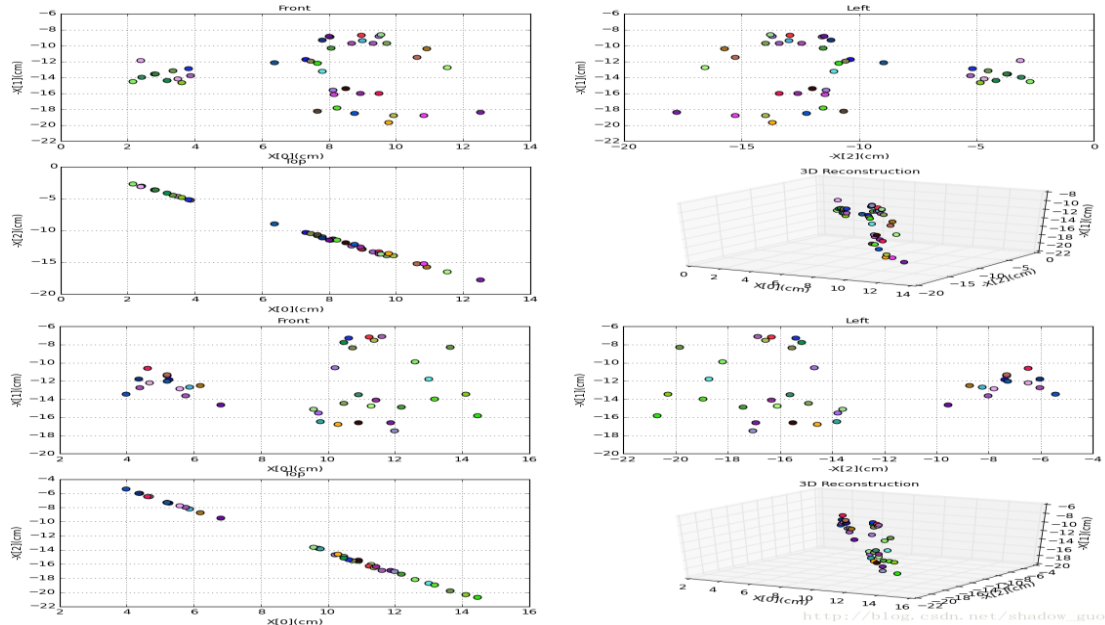


Fig6: Experiment Result 2.2

Improved reconstruction results. As shown in the figure, the set of top view feature points is more and more close to a straight line, and the distribution is more realistic than using only the sift method. The errors are 15.65128912 and 9.086802621 pixels. However, in the improved method, the more realistic affine structure we get will also reduce the error of reprojection, so it is reasonable to use the SIFT+KLT method to improve the consistency of the feature point set and the authenticity of the restoration. , can reduce the error of reprojection.

6. CONCLUSION

In this paper, the feature matching algorithm with the most information points and the highest accuracy is selected in the 3D modeling algorithm provided by the VTK class library. The KLT feature tracking is used to analyze and process the image, and the LK algorithm is analyzed in each frame of the scene. Calculate the change of direction of motion, try to combine the two algorithms and give a single sift algorithm and the reconstruction and data analysis of the two algorithms at the same time. Combine the two-dimensional and three-

dimensional algorithms with the highly applicable algorithm process. And give further optimization directions.

7. REFERENCES

- [1] Jiang Huaqiang, Cai Yong, Zhang Jiansheng, Li Zisheng. Research on 3D Reconstruction Algorithms Based on Improved SFM index of [J].Computer technology and application, 2019,45(2):88-92
- [2] Zhang Yanbo, Zhang Guofeng, Tian Baozhu. Analysis of sensibility and landslide control policy on factors of slope stability in opencast coal mine index of [J].Coal engineering, 2011, 1(5): 105-107.
- [3] Wang Kun, Zheng Nanning. 3D Face Modeling Based on SFM Algorithm index of [J].Journal of Computer Science, 2005,1(6):1049-1053.
- [4] Xue Junpeng, Su Xianyu, Xiao Yongliang, Liu Xiaoqing. A New BA Method Based on Binocular Vision. optoelectron-laser, 2011(06):889-892.