A Kind of new parallax algorithm based on symmetric continuous optimization

Kaiwen Jiang 1,2*, Degan Zhang 1,2, Haixia Xu 3

1,2Key Laboratory of Computer Vision and System (Tianjin University of Technology), Ministry of Education, 300384, China
1,2Tianjin Key Lab of Intelligent Computing & Novel software Technology, Tianjin University of Technology, Tianjin, China
3School of Electronic and Information Engineering, Tianjin Vocational Institute, Tianjin 300410, China
*Correspondence: jiang512@email.tjut.edu.cn

Abstract: With the fast development of 3D imaginations becomes more and more fascination, multi-view stereo based 3D reconstruction is a significant technique for those application. To facilitate the subsequent processing of 3D reconstruction and reduce the possibility of other algorithms falling into local optimal solutions, attempting to get better and faster performance, a new parallax calculation based on symmetric continuous optimization is proposed in this paper. The algorithm is proposed here will be tested respectively in the same data, in order to certific the algorithm applied is better than traditional minimize $E_1$ algorithm.

1. Introduction
Binocular stereo matching to calculate depth information is suggested in the depth calculation process, based on the multi-view stereo matching algorithm[1] of depth fusion. The relationship between the illustration and the depth map is reciprocal relationship. Therefore, through the calculation of the inspection map the depth map can be obtained easily. In this paper, the binocular stereo matching technology based on bandit energy function[2]. One of the main problems of binocular stereo matching based on non-convex energy function is how to make the solution closer to the global optimal solution[3], that is, to reduce the solution function into a local optimal solution. The main purpose of these paper, it's to build a new energy function and solving algorithms to reduce the possibility of the algorithm falling into a local optimum, improve the quality of inspection, and facilitate the subsequent processing of 3D reconstruction[4].

2. Principle of our algorithm
It is supposed that the images which need to be subjected to binocular stereo matching calculation have undergone polar correction, and the horizontal lines of the two images are aligned[5].It is assumed that the grayscale images $P_l$ and $P_r$ respectively represent the reference image (left image) and the target image (right image). At the same time, $d_l$ and $d_r$ respectively represent the corresponding left and right parallax. In the view of continuity: $P_l$ and $P_r$, $d_l$and $d_r$ are all continuous or piecewise continuous functions. Taking the left image and the left parallax as an example, the definition of the left image is $P_l: \Omega \rightarrow [0,1], (x,y) \rightarrow P_l(x,y)$. The left parallax is
defined as \( d_1 : \Omega \rightarrow R, (x, y) \rightarrow d_1(x, y) \). The point \((x, y)\) in the left image corresponds to \((x + d_1(x, y), y)\) in the right image. The purpose of the current calculation is to find the parallax \(d\) and \(d_r\) by given \(P_l\) and \(P_r\). From the perspective of probability, it is to calculate the parallax function corresponding to the maximum probability \(\sum P(d_1, d_r|P_l, P_r)\). According to the Bayes formula and the definition of conditional probability:

\[
P(d_1, d_r|P_l, P_r) \propto P(P_l, P_r|d_1, d_r) \cdot \sqrt{P(d_r|d_1)} \cdot \sqrt{P(d_1|d_r)}
\]

Maximizing the probability is equivalent to minimizing the entropy function:

\[
E(d_1, d_r|P_l, P_r) = -\ln P(P_l, P_r|d_1, d_r) - \ln \sqrt{P(d_r|d_1)} - \ln \sqrt{P(d_1|d_r)} P(d_r|d_1)
\]

Here, the entropy function is the ratio of the left and right parallax to the gray image. Assuming the parallax prior distribution, the likelihood distribution and the conditional probability distribution get the following forms:

\[
\begin{align*}
P(d_1) &= \frac{1}{W_1} e^{-\int \beta f(d_1)} \\
P(P_l, P_r|d_1, d_r) &= \frac{1}{W_2} e^{-\int \rho_{lr} - \int \rho_{rl}} \\
P(d_r|d_1) &= \frac{1}{W_3} e^{-\int \sigma((d_r|d_1))}
\end{align*}
\]

Mid-term, \(W_1, W_2, W_3\) are the normalized coefficients. Under the assumption of Lambertian surface, in different images, the same point in space has the same brightness. Therefore, a natural data item is defined as the pixel brightness difference, that is, if the parallax value is correct, the value of \(|P_l(x, y) - P_r(x + d_1(x, y), y)|\) should be close to 0. In the first few scenes, due to changes in ambient light or other factors, in different images, the same point in the space corresponding to the pixel bright spots may have slight changes. Hence, it cannot make the good robustness algorithm while only the brightness consistency. In order to improve the adaptability of the algorithm to the environment, gradient consistency \(|\nabla d_1 - \nabla d_r|\) introduces, to reduced the influence of noise and ambient light changes to the algorithm.

### 3. A new energy function

\(S_l\) and \(S_r\) are introduced in the reference image space and the target image space respectively. In fact, the resolutions of the images involved in the calculation are the same. Therefore, \(S_l\) equal to \(S_r\). If it does not cause confusion, \(S\) is adopted to represent the image space.

Minimization of the energy function: The energy function involves two functions \(d_1\) and \(d_r\), which are divided into two guards, and iterative calculations are performed through alternate solutions.

\[
\begin{align*}
E_l &= \iint_S |P_l(x, y) - P_r(x + d_1(x, y))| + \lambda |\nabla P_l(x, y) - \nabla P_r(x + d_1(x, y), y)| dxdy \\
&\quad + \iint_S \sqrt{|d_1(x, y)| - |d_r(x + d_1(x, y), y)|^2} dxdy + \iint_S |\nabla d_1(x, y)| dxdy
\end{align*}
\]

\[
\begin{align*}
E_r &= \iint_S |P_r(x, y) - P_l(x + d_r, y)| + \lambda |\nabla P_l(x, y) - \nabla P_r(x + d_l(x, y), y)| dxdy \\
&\quad + \iint_S \sqrt{|d_r(x, y)| - |d_l(x + d_r(x, y), y)|^2} dxdy + \iint_S |\nabla d_r(x, y)| dxdy
\end{align*}
\]

Bring the formula into Euler-Lagrangian theorem and minimize the formula. The formula is linearized by a fixed-point algorithm, and a fixed-point algorithm is used to iterate k+1 times of the left parallax \(d_1\), the initial value of the iteration is set to 0, and the nonlinear term in the formula is Taylor expanded to eliminate part of the nonlinearity.
\[ \text{div} \mathbf{\Phi} \left( |\nabla d_i^k + \nabla d_i^{k+1}|^2 \cdot (\nabla d_i^k + \nabla d_i^{k+1}) \right) = 0 \] (6)

Figure 1. The minimization process of Energy functional

Figure 2. The proposed algorithm flow chart
Next, the Jacobi iterative algorithm is used to solve the above linear function. Two problems will be encountered in the calculation: one is the problem of computational efficiency, and the other is the problem of falling into a local optimal solution.

In order to solve the above problems, by constructing a three-layer tower the possibility that the energy function is selected into the local optimal solution is reduced. And a smooth approximation of the high-resolution image can be gotten by down sampling. To avoid falling into local optimal solutions in these regions, the high-frequency components of the image are omitted. Perform energy functional calculation on low resolution to obtain a smooth approximation on high-resolution image and then project the parallax function into the high-resolution image to optimize. The result parallax of up calculation is further iterated. Through such a calculation method can prevent the algorithm from falling into the local optimal solution in some areas.

4. Experimental results and Comparison analysis

The algorithm is proposed here will be tested respectively in the same data. In order to certificate the algorithm applied is better than traditional minimize $E_I$ algorithm, when the compare is applied between the results of traditional minimize $E_I$ algorithm and the algorithm is proposed. The matlab software development tool in operating system of Windows 10 is the software experimental platform. PC is Lenovo of Intel I5 CPU, and memory is 4.00GB. In order to verify the algorithm, this paper uses the binocular stereo matching test image collection on Middlebury. The test image collection on Middlebury has four groups, namely: Tsukuba, Venus, Teddy and Cones.

(a) Calculation results of this algorithm  (b) Calculation result of Minimize $E_I$

Figure 3. The disparity calculated by the two algorithms on the Venus

(a) Calculation results of this algorithm  (b) Calculation result of Minimize $E_I$

Figure 4. The disparity calculated by the two algorithms on the Teddy
Figure 5. The disparity calculated by the two algorithms on the Cones
(a) Calculation results of this algorithm  (b) Calculation result of Minimize $E_1$

Figure 6. The disparity calculated by the two algorithms on the Tsukuba
(a) Calculation results of this algorithm  (b) Calculation result of Minimize $E_1$

Table 1. Quantitative evaluation results of the two models on the image set

|                | All(%) | Nonocc(%) | Disc(%) |
|----------------|--------|-----------|---------|
| Venus          |        |           |         |
| Minimize $E_1$ | 12.1   | 9.6       | 38.5    |
| Proposed Algorithm | 6.2   | 4.2       | 34.6    |
| Teddy          |        |           |         |
| Minimize $E_1$ | 28.9   | 20.9      | 43.2    |
| Proposed Algorithm | 24.9  | 16.4      | 37.5    |
| Cones          |        |           |         |
| Minimize $E_1$ | 27.2   | 18.6      | 30.9    |
| Proposed Algorithm | 20.6  | 10.9      | 25.6    |
| Tsukuba        |        |           |         |
| Minimize $E_1$ | 8.85   | 6.99      | 31.0    |
| Proposed Algorithm | 9.52  | 7.31      | 32.6    |

From the above analysis, it can be seen that the algorithm in this paper effectively improves the traditional continuous optimal parallax calculation by the results on the test images of Venus, Teddy, Cones, and Tsukuba. From the parallax map calculated by the two algorithms, especially the part in the red box, it can be clearly seen that the algorithm in this article effectively improves the quality of the parallax map. The parallax is closely related to the depth[6]. In the 3D reconstruction, it is also Will effectively improve the quality of the 3D reconstruction model.

References
[1] J. Bloomenthal. (1988) Polygonization of implicit surfaces[J]. Computer Aided Geom. Des. 1988,5(4): 341-355.
[2] Zhang DG, Kang XJ. (2012) A novel image de-noising method based on spherical coordinates system[J], EURASIP Journal on Advances in Signal Processing, (110):1-10 DOI:10.1186/1687-6180-2012-110.
[3] Jiang, KW, Zhang DG, Xu HX. (2019) Semi-supervised classification method of SAR images using spectral clustering in Contourlet domain [C]. 4th International Seminar on Computer Technology, Mechanical and Electrical Engineering, ISCME 2019-3:490-496

[4] Wang Z, (2010) Facial expression recognition based orthogonal supervised spectral discriminant analysis [C]. IEEE 10th International Conference on Signal Processing, 2010: p.1056-1059.

[5] W. Lorensen, H. Cline. (1987) Marching cubes: A high resolution 3D surface construction algorithm [C]. SIGGRAPH, 1987:163 ─ 169.

[6] G. Alberti, G. Bouchitte, G. Maso. (2003) The calibration method for the Mumford-Shah functional and free-discontinuity problems [J]. CalcNarTarital Differential Equations.2003.16(3):299-333.