Experiment of Stereo Matching Algorithm Based on Binocular Vision

Dongdong Lv$^1$ and Guie Jiao$^2,*$

$^1$Shanghai Ocean University, Shanghai, 201306
$^2$Shanghai JianQiao University, Shanghai, 201306

*Corresponding author: jiaoguie@126.com

Abstract. Binocular stereo matching is a hot topic in 3D reconstruction, target detection and target tracking. Through experiments, this paper analyzes and compares several classic binocular stereo matching algorithms, and finally selects SGBM with good comprehensive performance. Moreover, based on this, it gets accurate window size parameters through experiments, which provides a basis for the following 3D modeling.

Keywords: Error Percentage, Binocular Stereo Matching, Time Complexity, SAD Window

1 Introduction
About 80 percent of the information around us is obtained through visual perception[1]. As an important branch of computer vision, binocular stereo vision simulates human visual system to process real life. Since the 1960s, many of its achievements in scientific research have been applied to various fields and have affected every aspect of our lives. In 1982, Marry proposed the theory of computer vision and introduced theories and methods of visual algorithm in detail[2]. He thought the restoration of the geometric structure of a two-dimensional image to a three-dimensional image could be completed by algorithm. Accordingly, he proposed a complete set of algorithm theories and methods. At the end of the 20th century, Roy[3] believed that map cutting could effectively overcome the problem of algorithm optimization in stereo matching, and successfully verified his guessing through experiments. The key step of 3D modeling is stereo matching. Its basic step is to obtain images of the scene, and then use the corresponding algorithm to match the corresponding pixel. After successful matching, the disparity map and depth can be calculated[4].

In recent years, the application and research of 3D reconstruction in images or videos are quite popular, especially in the aspect of images. Stereo matching is the key to 3D reconstruction. This paper mainly introduces each link of 3D modeling and focuses on analyzing and comparing stereo matching. Among them, BM and SAD (Sum of absolute differences) algorithms[5][6] are local stereo matching, and SGBM algorithm[7] is a semi-global stereo matching algorithm.

2 The Principle of Binocular Vision
When the left and right cameras collect information on the same target, two different images can be
obtained. The difference between their positions is called disparity[8]. The measuring principle of binocular stereo vision is similar to that of human eyes, both of which are based on disparity principle, as shown in Fig. 1.

![Fig. 1 Principle of Binocular Stereo Imaging](image)

The optical centers of the two cameras are like human eyes. The distance between them is recorded as baseline B; the focal length of the camera is f. The coordinate of P at any point in space is recorded as $P(x, y, z)$. The two cameras observe the point from different angles and acquire the image of the feature point P from the left and right cameras. Suppose their plane coordinates are: $p_{sel} = (x_{sel}, y_{sel})$ and $p_{sw} = (x_{sw}, y_{sw})$.

Since the optical centers of the two cameras are flush, namely, the images of the two are on the same horizontal plane, we can know that the ordinate $y$ of the feature point P is the same. That is to say $y_{sel} = y_{sw} = Y$. According to the knowledge of triangulation, we can get it in formula (1).

$$
\begin{align*}
X_{left} &= f \frac{x}{z_c} \\
X_{right} &= f \frac{(x - B)}{z_c} \\
Y &= f \frac{Y_c}{z_c}
\end{align*}
$$

(1)

$$
\begin{align*}
x_c &= \frac{B X_{left}}{Disparity} \\
y_c &= \frac{B Y}{Disparity} \\
z_c &= \frac{B f}{Disparity}
\end{align*}
$$

(2)

Then, The disparity of the corresponding points of the two cameras’ images is: $Disparity = X_{sel} - X_{sw}$.

The 3D coordinates of the feature point P in the camera coordinate system can be calculated as in formula (2).

Thus, according to the above formula, we can use binocular camera to calculate the 3D coordinates of the same object or feature point.

3 The Basic Process of 3D Reconstruction

(1) Using binocular camera to collect two images with different angles;
(2) Camera calibrating to calculate camera internal and external parameters;
(3) Image processing and image feature extracting.

Finally, appropriate methods are used to match the corresponding points of the two images to calculate the positional deviation between the corresponding points to obtain the disparity map. This allows us to determine the depth image, recover the 3D information of the scene, and model it [9].

4 Stereo Matching Algorithm

4.1 BM Algorithm

BM algorithm is a local stereo matching one. The algorithm takes the support window as the theoretical basis and assumes feature vector traversal search in it. Then, it calculates the similarity between different windows and this window in the traversal, and takes the most similar window as the final result. The first task of this algorithm is to calculate the matching cost [10], so as to determine the
best search window.

4.2 SAD Algorithm
The absolute value of gray difference and SAD is a local image block matching algorithm. The basic principle is that the absolute value of the difference between the pixel value of the matching feature point and the corresponding pixel value in another image is summed to evaluate whether the two images are similar.

4.3 SGBM Algorithm
The local matching algorithm is fast and simple, but the disparity map obtained by experiments is rough and the algorithm is not robust. It is easily affected by external factors like illumination and occlusion. The global matching algorithm has high precision, so its complexity is also high. SGBM algorithm is a semi - global stereo matching algorithm. By selecting the difference of each pixel, SGBM forms a disparity map and sets a global energy function related to it, so as to minimize the energy function. Then, we can get the optimal gap of each pixel. It is as in formula (3).

$$E(d) = \sum_{p} C(p, d_p) + \sum_{q \in N_p} P_1 T[1|d_p - d_q| = 1] + \sum_{q \in N_p} P_2 T[|d_p - d_q| > 1] \quad (3)$$

In the formula, $\sum_{p} C(p, d_p)$ is the sum of all pixel matching costs based on the disparity $d$, which represents the data item; The latter two terms are the smoothing terms, $P_1$ and $P_2$ are the penalty coefficients of pixel $p$ in domain $N_p$.

5 Experimental Results and Analysis

5.1 Stereo Matching Effect of BM, SAD and SGBM Algorithms
The following uses indoor objects to correct the left and right images; since calculating the images’ PBM(total mismatch pixel percentage) requires the standard disparity map of the test image, the experiment uses the original images and standard disparity map of the authoritative website of the Middlebury gallery, which is specially provided for stereo matching algorithms. Left view is as shown in Fig. 2, right view is in Fig. 3.

![Fig. 2 Left View](image1)
![Fig. 3 Right View](image2)

Matching effects are shown is Fig4. Among them, Fig. a is BM disparity map, Fig. b is BM standard disparity map, Fig. c is SAD disparity map, Fig.d is SAD standard disparity map, Fig. e is SGBM disparity map, Fig. f is SGBM standard disparity map. In the algorithm running, the running time of BM is 7ms, SAD’s is 4000ms, SGBM’s is 155ms.
Fig. 4 Running Results of Three Algorithms

The disparity map obtained from the experiments is analyzed and compared with the standard disparity map in Middlebury, and then the accuracy of each algorithm is analyzed and judged. The commonly used criteria are algorithm running time and mismatched pixel percentage (PBM). The calculation formula of mismatched pixel percentage is shown in the following formula (4):

\[ B = \frac{1}{N} \sum_{(x, y)} |d_c(x,y) - d_f(x,y)| > \varphi \] (4)

In the above formula, N represents the number of pixels in the entire image. This index reflects the ratio of the pixels calculated by the algorithm that the error between the value of the disparity and the value in the standard disparity map is greater than a certain threshold among the pixels of the entire image. The threshold should not be too big, otherwise the data obtained lacks credibility.

5.2 PBM Values of Each Algorithm with Different \( \varphi \) Values

In formula 4, \( \varphi \) value has an effect on the experimental results of BM and SGBM. The 20 sets of data are collected to select \( \varphi \) value, the following 12 sets of data is as shown in Table 1:

In the above parameters, when \( \varphi \) value is between 1 and 4, the ratio of mismatched pixels of BM algorithm is better than that of SGBM, while that of SGBM algorithm is relatively high and fluctuate little. For the intuitiveness and effectiveness of the algorithm comparison, we use \( \varphi = 5 \), the experimental results are shown in Fig. 5, among them, Fig. a is BM error map, Fig. b is BM pseudo-color error map, Fig. c is SAD error map, Fig. d is SAD pseudo-color error map, Fig. e is SGBM error map, Fig. f is SGBM pseudo-color disparity map.

In the above experimental results, in order to see the error effect more intuitively, the pseudo-color error maps are attached. In the pseudo-color maps, black indicates that there is no error, and where there is color, there is an error. Different colors indicate different degrees of errors.

Table 1. PBM Value of Each Algorithm with Different \( \varphi \) Values

| Algorithm \( \varphi \) | SGBM | BM | SAD |
|-----------------------|------|----|-----|
| 1                     | 0.5434 | 0.3128 | 0.9935 |
| 2                     | 0.3796 | 0.2568 | 0.9864 |
| 3                     | 0.2909 | 0.2447 | 0.9718 |
| 4                     | 0.2389 | 0.2389 | 0.9604 |
| 5                     | 0.1986 | 0.2342 | 0.9507 |
| 6                     | 0.1727 | 0.2298 | 0.9411 |
| 7                     | 0.1545 | 0.2255 | 0.9305 |
| 8                     | 0.1347 | 0.2221 | 0.9224 |
| 9                     | 0.1214 | 0.2179 | 0.9137 |
| 10                    | 0.1111 | 0.2121 | 0.9086 |
| 11                    | 0.1033 | 0.2081 | 0.9051 |
| 12                    | 0.0972 | 0.2050 | 0.9023 |
Basics modeling optimal. It still is algorithm, through sizes. Affect algorithm, later is acceptable. Its effect. Moreover, when the SAD window of SGBM is set to a value around 10, the error is optimal. Later research will also apply this algorithm to 3D modeling of furniture so as to show the modeling effect.

5.3 The Selection of SGBM Algorithm Parameters
In SGBM algorithm, the main factor that affects SGBM is the SAD window size, which will directly affect the SGBM matching effect. The relative data show different results for different SAD window sizes. Due to the limited space, the experimental data and the experimental results are not given here. Through many experiments and analysis, when the SAD window size is around 10, the SGBM error is optimal.

Conclusion
It can be seen from the above table that BM algorithm runs the fastest, but the mismatched pixel value is very big, and the matching effect is not good. From the experimental results, compared with BM algorithm, the PBM value and running time of SAD algorithm are greater than BM, and this algorithm is not stable. Although the running time of SGBM algorithm is longer than that of BM algorithm, it is still acceptable. Its advantage is that the mismatched pixels are relatively low and the matching accuracy is high. Moreover, when the SAD window of SGBM is set to a value around 10, the error is optimal.

References
[1] Ma Songde, Zhang Zhengyou. Computer Vision—Basics of Computing Theory and Algorithms[M]. Beijing: Science Press, 1997.
[2] Marr DC. A Computational Investigation into the Human Representation and Processing of Visual Information [M]. San Francisco: W. H. Freeman and company, 1982.
[3] Bao Zongfan, Li Honghua. Improvement of Laplace Operator on Edge Detection[J]. Journal of...
China Institute of Metrology, 2000(02):81-84.

[4] Wang Jue, Zeng Jianping, Zhou Baohua, etc. Online Forum Opinion Leaders Discovering Method Based on Clustering Analysis[J]. Computer Engineering, 2011,37(05): 44-46+49.

[5] He Renjie. Improvement of Regional Related Match Algorithm for Binocular Stereo Vision and Its Implementation[J]. Modern Electronics Technique, 2009,32(12):68-70.

[6] Sun Daohui, Sun Lijun, Chen Tianfei. Binocular Stereo Matching Approach based on OpenCV[J]. Software Guide, 2018,17(09):92-94+100.

[7] Wang Xinyan, Pan Wei, Wang Yuelian, Liu Xinyue. Stereo Matching Algorithm Based on Global Error Energy Function[J]. Computer Engineering, 2017,43(07):244-249.

[8] Zhou Fang. Research and Implementation of Stereo Matching Algorithm in Binocular Vision[D]. Dalian University of Technology, 2013.

[9] NIU, Zhigang, LI, Li jun, Wang, Tie, 3D Reconstruction Based on Binocular Stereo Vision of Robot[J]. AdvMater Res, 2011,338: 645-648

[10] Ning Xiaofei. Research on Semi-Global Stereo Matching Algorithm in Binocular Stereo Vision[D]. Liaoning University, 2014.