Stereo Matching Based on Visual Sensitive Information

Hewei Wang*,†, Muhammad Salman Pathan†‡, and Soumyabrata Dev†‡ §

* Beijing University of Technology, Beijing, China
† School of Computer Science, University College Dublin, Ireland
‡ ADAPT SFI Research Centre, Dublin, Ireland
§ Beijing-Dublin International College, Beijing, China
Problem statement

- Problem statement
- Related works
- Contribution of this paper
The image source of this technology is usually captured by a binocular camera, which can provide the left and right perspectives of the image, to carry out the stereo matching operations.
**Related work**

Wang et al. constructed a real-time matching system, which can get a disparity image quickly.

Jian Sun’s work, a single direction stereo matching algorithm is proposed, which limits the constraints of matching between two images and finds the optimal pixels.

Ysioncenelli et al. proposed a very classic Sum of Absolute Differences (SAD) stereo matching algorithm based on the difference between two gray regions.

Savoy et al. used ground-based sky cameras in a stereo setup to compute the cloud-base height.

Yin Chuanli et al. segmented the image by color threshold, fitted the segmented region, and optimized the parallax results by minimizing the energy function.

Using the similar method, Da et al. furthermore improved the algorithm and used the gray features of the image to find the optimal matching region.
Contributions of this paper

1. We provide a systematic analysis of the related frame works on stereo matching and identify the key shortcomings in them.

2. A novel stereo matching algorithm is proposed wherein we introduce the visual sensitivity factor in the cost calculation stage of the matching process to improve the matching degree.

3. All the codes related to the different experiments in the paper are open-source. The proposed model code, dataset, and experimental results are available at https://github.com/WangHewei16/Stereo-Matching
Proposed Method

- Overall Diagram
- Analysis of Novel Algorithm
Analysis of Novel Algorithm

$H_x$ is the horizontal convolution template whereas the $H_y$ is the vertical convolution template. $G_x$ is the transverse brightness value and $G_y$ is the longitudinal brightness value. We can **get the brightness difference value** from these template and formula.

$$G_x = H_x \ast A$$

$$G_y = H_y \ast A$$
Taking $G$ as the second generation value to guide a certain region matching between two images, the first cost is to use the binary string generated according to the gray difference whose hamming distance represents the gray approximation of the two regions.

\[
G = \sqrt{G_x^2 + G_y^2}
\]

The process of using two-generation values to judge in turn is represented by the following formula:

\[
f_s(p, d) = 1 - \exp\left[-\frac{C_s(p, d)}{\lambda_s}\right]
\]

\[
\Rightarrow f_G(p, d) = 1 - \exp\left[-\frac{C_g(p, d)}{\lambda_g}\right]
\]
Analysis of Novel Algorithm

According to the calculation method of the second generation value mentioned above, we can easily get the direction of visual sensitivity in a certain area. The calculation is as follows:

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

This angle is called as the visual sensitivity factor of a target pixel. The visual sensitivity direction of different position points represents the direction of the fastest gray change, and each point has such a direct factor.
Four degrees of freedom growth decision conditions

\[ D_s(p_1, p) < L_1 \]

\[ D_c(p_1, p) < f(\tau_1, \theta) \]

\[ D_c(p_1, p_1 + (1, 0)) < \tau_2 \]

\[ D_c(p_1, p) < f(\tau_3, \theta), \text{ if } L_1 < D_3(p_1, p) < L_2 \]

The precondition of the four degrees of freedom growth decision conditions

\[ (\tau_3 < \tau_2 \leq \tau_1, L_1 < L_2) \]
Experiment and Discussion

- Dataset
- Subjective evaluation
- Objective evaluation
- Impact of noise on mismatch rate
Middlebury official test dataset is utilized in this paper for the experimentation. This dataset have also used in other many literature. The dataset mainly used for computer vision processing of binocular images including image restoration, stereo matching, image recognition and other. Because of space constraint, this paper mainly selects four kind of images from Middlebury 2.0 for experiments containing 24 left and right view color images and the real parallax images.
This paper presents a subjective form to show the image matching effect and compares the test image obtained by proposed algorithm with the test image obtained by the classical algorithm.

The first column is the perspective image of the original image from the experimental data set, the second column is the ideal parallax image for reference, the third and fourth columns are the resultant images obtained by the classical algorithm census, and the fifth and sixth columns are the images obtained by stereo matching algorithms based on visual sensitive information. Among them, the red marks in the fourth and sixth columns indicate the mismatching areas when comparing the two resultant maps with the ideal disparity map.
**Method:** We test the error matching rate through the four types of images processed by the proposed algorithm and compares with other matching algorithms of the same kind.

**Result:** Table I that the error matching rate of the proposed algorithm is lower than that of the same kind of algorithm, and the overall matching effect is better.

**TABLE I: The error matching rate obtained by various algorithms**

| Algorithm        | cones | tsukuba | teddy | venus  | avg%   |
|------------------|-------|---------|-------|--------|--------|
| Census           | 21.89 | 27.41   | 23.43 | 27.78  | 25.13  |
| Proposed         | 4.10  | 2.01    | 7.84  | 1.35   | 3.82   |
| Reference [23]   | 4.04  | 2.53    | 7.57  | 1.60   | 3.93   |
| SG-C [24]        | 12.92 | 4.80    | 8.05  | 1.91   | 6.92   |
| Mp-C [22]        | 6.78  | 4.50    | 11.32 | 3.55   | 6.53   |
Method: To prove this, we added different concentrations of Gaussian noise to the above four original images, and it is mainly expressed by the change of average mismatch rate with the increase of noise concentration.

Results: The table proves that when the noise concentration increases, the change in the degree of error matching rate and the sensitivity to noise remains low.

| Noise     | none  | 2%    | 5%     | 10%    | 15%    |
|-----------|-------|-------|--------|--------|--------|
| Ref. [25] | 4.54  | 6.67  | 11.73  | 26.88  | 48.63  |
| Proposed  | 3.82  | 5.02  | 8.65   | 17.33  | 38.46  |

TABLE II: Mismatch rate under different Gaussian noise concentrations
Conclusion
Conclusion

Summary of content

- The stereo matching technology is studied.
- A stereo matching algorithm based on visual sensitive information is proposed.
- In the cost calculation stage, the sensitive information factor is proposed, and the robustness of cost matching is enhanced.
- In the process of aggregation, the aggregation template is extended for each pixel according to the sensitivity factor.
- In the experimental part, the proposed algorithm is demonstrated subjectively and proved by objective data through the combination of graph and tables.
- Summing up, the proposed algorithm can effectively match the left and right view of images, and the matching effect is better than a variety of similar algorithms.

Future work

- We intend to further study the factors that have a direct impact on the parallax value, and the actual scene image depth map generation and three-dimensional environment reconstruction work, as far as possible to capture more actual scene images to form their data set.