Research on new stereo matching algorithm of multi-rotor UAV based on vision


Published in: Advances in Manufacturing Technology XXXIII

Document Version: Peer reviewed version

Queen's University Belfast - Research Portal: Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2019 IOS. This work is made available online in accordance with the publisher’s policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 11. Jan. 2020
Research on new stereo matching algorithm of multi-rotor UAV based on vision

Haixin WANG a, Yan JIN b, Jianxin SHEN a*

a College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China.
b School of Mechanical and Aerospace Engineering, Queen’s University Belfast, Belfast, UK

Abstract. The application of multi-rotor UAV in vision has been developed rapidly in recent years, and multi-rotor unmanned aerial vehicles (UAV) can’t be widely used without the aid of visual system. The application of visual technology in UAV has attracted more and more attention with the continuous development of image processing and environment reconstruction technology. Getting accurate depth map from the stereo image has always been the pursuit of researchers. Stereo matching is the most important part of stereo vision. An improved matching algorithm is proposed to solve the problems of high image noise and low image accuracy in binocular vision imaging of multi-rotor UAV after binocular information fusion. Experimental results show that the proposed algorithm has advantages in accuracy compared with other algorithms. The validity and reliability of the proposed stereo matching algorithm for multi-rotor UAV based on binocular vision are verified by indoor simulation tests and real environment tests of a UAV.

Keywords. Matching algorithm, UAV, Binocular vision.

1. Introduction

UAV are relatively simple in construction and operation, and play an important role in aerial photography, rescue and search, agricultural plant protection, earthquake relief and other fields. The autonomous obstacle avoidance function of UAV can ensure the UAV to complete complex and difficult tasks. With the development of machine vision technology, binocular vision technology is gradually applied to the obstacle avoidance of UAV.

In terms of obstacle avoidance in machine vision, in 2013, Tomoyuki and Sebastian used monocular vision to collect continuous images in front[1]. SURF feature point extraction and matching algorithm were used to used to measure the expansion coefficient of obstacles. The coefficient was proportional to collision time. In 2016, Mohamed Elawady’s team used optical flow as the motion parallax of the UAV to calculate the optical flow diagram of two consecutive frames of images[2]. Pixel matching was carried out for two optical flow graphs, and the error matching points were removed according to the Euclidean distance of the matching
The image was vertically divided into five regions, and the average or median amplitude of non-discarding points in each region was calculated.

Aiming at this problem, this paper improves the weight allocation strategy of the traditional adaptive weight algorithm from the color feature and distance feature.

2. Principle of adaptive weight matching algorithm

In essence, the adaptive support weight algorithm proposed by Yoon assigns appropriate weights to each pixel in the window according to its color similarity and spatial proximity to the point to be matched. In general, the larger the geometric distance between the pixel q in the window and the pixel P to be matched, the greater the likelihood that they are not on the same surface. Meanwhile, the greater the color difference the color difference between them, the greater the probability that they do not belong to the same part of the scene. Therefore, when the color of any pixel in the window is closer to the pixel to be matched and the geometric distance is closer, a greater weight will be assigned to it.

The specific process is as follows: in window $\Omega_p$, the color difference between any pixel $q$ and pixel $p$ to be matched is denoted as $\Delta_c(p,q)$, and the geometric distance between two pixels is denoted as $\Delta_g(p,q)$. Since they can be considered as independent, the support weight $w(p,q)$ of pixel $q$ can be denoted as:

$$w(p,q) = f_c(\Delta_c(p,q)) \cdot f_p(\Delta_g(p,q)), \quad (1)$$

In formula 1, color similarity intensity

$$f_c(\Delta_c(p,q)) = \exp\left(-\frac{\Delta_c(p,q)}{\gamma_c}\right) \quad (2)$$

Where $\Delta_c(p,q)$ is the color similarity of window pixel $q$ and pixel $P$ to be matched, which is represented by the Euclidean distance between them in the colour space of Lab. $L$ stands for brightness in Lab space; A positive is red, negative is green; The positive number of $b$ is yellow, which can be converted from RGB space. Let the coordinates of pixel $P$ and $Q$ in the Lab space are $c_p = [L_p, a_p, b_p]$ and $c_q = [L_q, a_q, b_q]$, then $\Delta_c(p,q)$ can be expressed as

$$\Delta_c(p,q) = \sqrt{(L_p - L_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2} \quad (3)$$

In formula (1), the spatial proximity intensity
\[ f_p(\Delta_g(p,q)) = \exp \left( -\frac{\Delta_g(p,q)}{\gamma_g} \right) \]  

(4)

Where, \( \Delta_g(p,q) \) represents the geometric proximity between pixel points \( p \) and \( q \), which is the Euclidean distance between them in the image plane. If the coordinate of pixel \( q \) is \( q(x_q, y_q) \) and the coordinate of pixel \( p \) is \( p(x_p, y_p) \), then

\[ \Delta_g(p,q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \]  

(5)

According to equations (2) to (4), the support weight of pixels:

\[ w(p,q) = \exp \left( -\frac{\Delta_c(p,q)}{\gamma_c} - \frac{\Delta_g(p,q)}{\gamma_g} \right) \]  

(6)

Among them, \( \gamma_c \) and \( \gamma_g \) are used to adjust the influence of color similarity and distance proximity on support weight, which is selected according to experience.

\[ C(p, \bar{p}) = \frac{\sum_{q \in \Omega_p, q \neq \bar{p}} w(p,q)w(\bar{p},q)\delta(q,\bar{q})}{\sum_{q \in \Omega_p, q \neq \bar{p}} w(p,q)w(\bar{p},q)} \]  

(7)

\[ \delta(q,\bar{q}) = \min \left\{ \sum_{c \in \{r, g, b\}} \left| I_c(q) - I_c(\bar{q}) \right|, T \right\} \]  

(8)

3. Improved adaptive support weight algorithm

The perception of color also has two remarkable characteristics. First, the sensitivity to light intensity is higher than the sensitivity to color. Second, the feeling of color depends on the hue and saturation of color, and has nothing to do with light intensity. In order to facilitate color processing and recognition, human visual system often adopts HSI color space, which is widely used in image processing and computer vision.
The HSI model is a double-cone coordinate system based on cylindrical polar coordinates, as shown in figure 1. Where, the hue component $H$ is the Angle between the color vector and the red vector, the saturation component $S$ is the length of the color vector, and the intensity component $I$ is the distance from the vertex of the lower cone to the color component.

**Tonal component:**

$$H = \begin{cases} \theta, B \leq G \\ 360 - \theta, B > G \end{cases}$$  \hspace{1cm} (9)

Among

$$\theta = \arccos \left\{ \frac{0.5[(R - G) + (R + B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$  \hspace{1cm} (10)

**Saturation component:**

$$S = 1 - \frac{3}{R + G + B} \left[ \min(R, G, B) \right]$$  \hspace{1cm} (11)

**Intensity component:**

$$I = \frac{1}{3}(R + G + B)$$  \hspace{1cm} (12)

In the HSI color model, the value ranges of hue component $H$ and saturation component $S$ are both 0-1, the value range of the light intensity component is 0-255, in order to avoid the light intensity component on the determination of color similarity has too decisive impact.
In formula 2, the colour similarity intensity becomes

$$f'_s(\Delta_c(p,q)) = \exp\left(-\frac{\Delta'_c(p,q)}{\gamma_c}\right)$$  \hspace{1cm} (13)

The final weight expression is

$$w(p,q) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\Delta^2}{2\sigma^2\gamma_g} - \frac{\Delta'_c(p,q)}{\gamma_c}\right)$$  \hspace{1cm} (14)

4. Parallax optimization

Because of the different camera positions, the foreground has different occlusion to the background, so that one point in the left image doesn’t have a corresponding point in the right image. However, in the matching process, the matching algorithm will definitely give the parallax value, which will lead to a decrease in the matching accuracy. Therefore, the left-right consistency check (LRC) was used to detect the occlusion points in the image and the parallax value was filled in.

The practice of left-right consistency checking is as follows: First, the left and right parallax images are taken as reference images and target images respectively to obtain the left and right parallax images. For any pixel P in the left parallax graph, the parallax is denoted as \(d_l\), and the parallax value corresponding to pixel P in the right parallax graph is \(d_r\), and the threshold value \(\varepsilon\) is set (generally 0, 1, 2). If \(|d_l - d_r| \leq \varepsilon\), then P is a valid match point; otherwise, it is denoted as an invalid match point. When all invalid match points are removed, the parallax value is filled. For an invalid match point, the first valid match point is found to the left and right respectively. The parallax of the two points is denoted as \(d_l\) and \(d_r\), and the parallax of P is set as \(d_p = \min(d_l, d_r)\), that is to take the parallax of the rear view as the parallax of the occluded part.

5. Experimental results and analysis

In the experiment, Venus, Teddy and Cones were measured with standard stereo images on the Middlebury test platform. The computer hardware is equipped with Inter Core i7 CPU, the main frequency is 3.6ghz, and the memory capacity is 8GB. The software programming environment is Visual Studio 2010 and OpenCV2.4.

In order to verify the influence of the two weight assignment criteria on the algorithm performance, the algorithm using only gaussian distributed spatial proximity strenght and HSI spatial color similarity were tested respectively. The test results of PBM(percentage of bad matching) provided by the Middlebury platform were shown in table 1. Where “Non” “All” and “Disc” respectively represent the mismatching rate of non-occluded regions, the mismatching rate of all regions, and the mismatching rate of deeply discontinuous regions, and Avg is the average mismatching rate of the above three.
**Table 1.** Experimental result with only a weight allocation rule is improved

<table>
<thead>
<tr>
<th>Condition</th>
<th>Venus</th>
<th>Cones</th>
<th>Teddy</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non</td>
<td>All</td>
<td>Disc</td>
<td>Non</td>
<td>All</td>
</tr>
<tr>
<td>Gauss only HSI</td>
<td>0.58</td>
<td>0.87</td>
<td>4.12</td>
<td>3.45</td>
<td>9.53</td>
</tr>
<tr>
<td>AS only</td>
<td>0.66</td>
<td>1.05</td>
<td>5.26</td>
<td>3.74</td>
<td>9.77</td>
</tr>
<tr>
<td>W</td>
<td>0.71</td>
<td>1.19</td>
<td>6.13</td>
<td>3.97</td>
<td>9.79</td>
</tr>
</tbody>
</table>

**Table 2.** Comparison of improved algorithm and other algorithm

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Venus</th>
<th>Cones</th>
<th>Teddy</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non</td>
<td>All</td>
<td>Disc</td>
<td>Non</td>
<td>All</td>
</tr>
<tr>
<td>RegionTreeDP[^5]</td>
<td>0.22</td>
<td>0.57</td>
<td>1.93</td>
<td>6.31</td>
<td>11.9</td>
</tr>
<tr>
<td>DoubleBP[^6]</td>
<td>0.31</td>
<td>0.45</td>
<td>1.87</td>
<td>2.90</td>
<td>8.78</td>
</tr>
<tr>
<td>HEBF[^5]</td>
<td>0.22</td>
<td>0.33</td>
<td>2.41</td>
<td>2.78</td>
<td>9.28</td>
</tr>
<tr>
<td>Cross-Based[^6]</td>
<td>0.62</td>
<td>0.96</td>
<td>3.20</td>
<td>6.28</td>
<td>12.7</td>
</tr>
<tr>
<td>ASW</td>
<td>0.71</td>
<td>1.19</td>
<td>6.13</td>
<td>3.97</td>
<td>9.79</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.54</td>
<td>0.82</td>
<td>3.81</td>
<td>3.37</td>
<td>9.43</td>
</tr>
</tbody>
</table>

**6. Conclusion**

An improved adaptive weight algorithm is proposed and tested on the VS2010 platform. On the one hand, the weight distribution is changed to improve the matching accuracy by changing the distance weight calculation method in the method window. On the other hand, HSI colour model is used to calculate the colour similarity to further improve the accuracy of parallax calculation. The left and right consistency check and median filter are used to optimize the parallax graph. Finally, it is proved that the improved algorithm can effectively reduce the mismatching rate of the original algorithm in different regions, and the matching effect is better than some other local algorithm through the experimental comparison with other methods.

**Acknowledgements**

This research was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX18_0317).
References


