Optimization and Innovation of SIFT Feature Matching Algorithm in Static Image Stitching Scheme

Yuan Jiao*

Tianjin University of Commerce Boustead College, Tianjin 300384, China

*Corresponding author e-mail: jiaoyuanb@boustead.edu.cn

Abstract. With the development of computer technology, people can obtain electronic images by a variety of means, but these images often need to obtain a panoramic view of a large field of view from multiple limited field of view images because the field of view is smaller than that of humans, and image stitching technology can be very good to solve the contradiction between vision and resolution. This paper studies the extraction and registration of image feature points and its related technologies, including image feature detection and extraction, point feature-based matching, and image fusion. According to the requirements of matching and image splicing, the feature points with good robustness to image deformation and occlusion are used as image features to indicate that it meets the requirements of image splicing registration. This paper studies and designs an improved algorithm to improve the accuracy of feature points extraction. Further research and analysis of the SIFT operator based on the characteristics of the invariant technology realize that the feature points extracted from the image can be used for reliable matching of an object or scene from different perspectives.

Keywords: Electronic Image, Feature Point, Image Stitching, SIFT, Corner Feature

1. Introduction

With the development of computer technology, electronic images have become one of the important materials for image analysis and data processing. At present, electronic image sequence data with high resolution can be obtained through electronic digital imaging equipment. Since the field of view of the image taken by the electronic device is smaller than that of a human, it is often necessary for a professional to stitch two or more images together to obtain a panoramic view of a large field of view. Due to the inherent contradiction between the resolution and the field of view in the vision system, in order to ensure higher resolution, the scene area corresponding to a single electronic image is usually small and the scene range is limited, which is not conducive to accurate and comprehensive description and analysis. Scenes. The image splicing technology is an effective way to solve this problem. Through image splicing, a series of real scene images describing the world can be spliced into a larger, single, and more complex large-scale scene map, which is more accurate and comprehensive. The description of the real world.
2. Extraction Principle of Image Feature Points
The most common method in image synthesis is image registration. This method can improve the accuracy of matching while reducing the amount of calculation. The reduction of the amount of calculation and the elimination of noise are mainly achieved by selecting the key information in the image, and it can also adapt to various images. The gray level changes, graphics occlusion and deformation of the image [1]. Compared with other methods, the registration method based on feature points can tolerate larger image differences, which helps to obtain stable registration results. The advantage of the SIFT method lies in the extracted feature points [2]. On the one hand, it can be used for reliable matching of an object or scene from different perspectives, on the other hand, because it does not change with changes in image scale and rotation, And it is robust to light changes, noise, and affine changes. Therefore, they are all highly unique. In order to enhance the matching stability and improve the anti-noise ability, it is necessary to eliminate those low-contrast key points and unstable edge response points, and then combine the fitting three-dimensional quadratic function to accurately obtain the position and scale of the key points. Only in this way can the goal be achieved.

3. Optimization and Improvement Based on Gaussian Second-Order Differential Pair Matching Algorithm
After in-depth research on the SIFT algorithm feature point detection operator, an improvement of the SIFT feature matching algorithm based on the Gaussian second-order difference feature point operator is proposed. The Gaussian second-order difference pyramid detection can be used to replace the extreme points of the Gaussian difference pyramid. Detection, simplifying its structure and reducing its algorithm, not only improves the real-time performance of the SIFT feature matching algorithm, but also maintains its accuracy. Scale invariance is the most typical feature of the Gaussian difference algorithm. This theoretical basis is the rule space filtering and automatic scale selection feature detection theory. According to the plane geometry theory, the local extreme points of the original function correspond to the first derivative function of the function Is the point of zero. It can be seen that the zero-crossing point of the first-order derivative of the function is actually the point where the original scale normalized function obtains the local maximum or minimum value. Because the signal feature of "zero-crossing point" is better to detect and obvious than extreme points, the author believes that the zero-crossing point detection of the first derivative function of the scale normalization function can be used to replace the local extreme point detection of this function. Therefore, this paper proposes a Gaussian second-order difference feature detection algorithm. A difference operation is performed on the Gaussian difference function to obtain the Gaussian second-order difference function. The zero-crossing point detection of the Gaussian second-order difference function is used to replace the extreme point detection of the Gaussian difference function. To determine the scale and feature points, which is:

\[ D(x, y, k\sigma) - D(x, y, \sigma) = D^2(x, y, \sigma) \]  

Among them \(D^2(x, y, \sigma)\)It can represent the second-order difference function of Gaussian. because:

\[ \frac{\partial D}{\partial \sigma} = (D(x, y, k\sigma) - D(x, y, \sigma))/(k\sigma - \sigma) = (D^2(x, y, \sigma))/(k\sigma - \sigma) \]  

and so:

\[ D^2(x, y, \sigma) = D(x, y, k\sigma) - D(x, y, \sigma) = \frac{\partial D}{\partial \sigma}(k\sigma - \sigma), k\sigma - \sigma \neq 0\]  

\[ D^2(x, y, \sigma) = 0 \leftrightarrow \frac{\partial D}{\partial \sigma} = 0 \]  

It can be seen that the zero-crossing point of the second-order Gaussian difference function is the point where the first derivative of the Gaussian difference function is zero, which is the local extreme point of the Gaussian difference function. Therefore, the zero-crossing point of the Gaussian second-order difference function pyramid can be detected Local extreme point detection instead of Gaussian difference pyramid to realize feature point positioning.
ifL(x,y,σ) is the image scale space function, the Gaussian function acts as the convolution kernel, and the Gaussian function equation is shown in (4):

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-\left(x^2 + y^2\right)}{2\sigma^2}\right)$$ \hspace{1cm} (4)

IfI(x,y,σ) is the input image, the input image is convolved, and the equation is equation (5):

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y)$$ \hspace{1cm} (5)

Continue to apply the difference of the Gaussian function to the image convolution operation, and the stable feature points in the scale space can be effectively detected. At this time, the Gaussian difference function DOG [3] is expressed by equation (6):

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \otimes I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma)$$ \hspace{1cm} (6)

Get the most stable image feature points [4].

3.1 Determining the direction of feature points

It is stipulated in the image attributes that a direction must be assigned to each feature point. There are many ways to give direction to feature points. This article introduces a method that is the most stable to the experimental results. Each image L(x, y) is obtained in advance through the difference of the pixels, and its gradient value m(x, y) and direction B(x, y) are expressed as equations (7) and (8):

$$m(x, y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x+1,y) - L(x,y-1))^2}$$ \hspace{1cm} (7)

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right)$$ \hspace{1cm} (8)

In the whole process, the stability of feature point matching has been significantly improved, which is worthy of recognition. 3.2 Realize the acquisition experiment of feature points

Figure 1 and Figure 2 show the SIFT feature points detected in the figure. In this paper, arrows are used to represent SIFT features, and the position of the point pointed to by the arrow represents the two-dimensional coordinate position of the key point in the image. In this example, the ratio threshold is 0.4. Comparing the matched SIFT in the two images, it is found that the accuracy and stability of position information, directionality and scale information are all very high, which is very consistent with the intuitive judgment of the human eye. Compared with SIFT features, the range of structural information described is also wider, but SIFT matching takes into account the scale and direction of features, so the matching ability of wide baseline images with large geometric distortions will be greatly improved by this method [5-6].

![Figure 1. SIFT features of the extracted image](image1)

![Figure 2. SIFT features of the extracted image](image2)
4. Image Registration Based on Optimized SIFT Feature Points

Under the influence of various interference factors, we can usually extract 3 or more matching SIFT feature points. At this time, the value of the affine transformation matrix $T$ can be estimated according to the least square method. Let $n$ be used to estimate the transformation parameters. Matching the number of feature points, the corresponding error calculation formula is shown in (9):

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} \left\| P_i - TP_0 \right\|_2 / n
\]  

(9)

The curve fitting process refers to the above-mentioned estimation of the transformation matrix $T$ by the least square method. Generally speaking, the statistical reliability of the fitting result is determined by the number of sample points used in the fitting process. The more points, the higher the reliability.  

4.1 Image mosaic

Image mosaic is to combine the registered image with the existing mosaic result in space in a certain way, so that the panoramic image can get a good mosaic effect. It must be based on the image registration. In terms of visual characteristics, this article is bold and innovative. By improving the weighted average method in the literature [7], a new color fusion method is proposed, namely the Gauss model. This model not only takes into account the differences in the overlap between the registered image and the existing stitching results, but also meets the human visual characteristics, effectively solving the problem of false boundaries caused by image differences. At the same time, we define the Gauss function related to the geometric size of the overlapping area as the weighting coefficient of this method. The new fusion strategy is modified as shown in formula (10):

\[
I_s'(x, y) = w_1I_s(x, y) + w_2I_r(x, y)
\]  

(10)

Among them, $I_s'$ is the new stitching result, $I_s$ is the existing stitching result, $I_r$ is the registered image, $w_i (i = 1, 2)$ represents the weight, and their relationship must meet the following requirements:

- if $(x,y) \in I_s$ and $(x,y) \notin I_r$, so $w_1 = 1$, $w_2 = 0$
- if $(x,y) \notin I_s$ and $(x,y) \in I_r$, so $w_1 = 0$, $w_2 = 1$
- if $(x,y) \in (I_s \cap I_r)$, first calculate according to the following formula $w_i (i = 1, 2)$:

\[
w_i = \exp(-\frac{1}{2} \left( \frac{(x-x_i)^2}{\sigma_i^2} + \frac{(y-y_i)^2}{\sigma_i^2} \right))
\]  

As shown in formula (11),

\[
w_i = 1 - w_2
\]  

(11)

As shown in formula (11),

\[
\begin{align*}
G_s(x, y) - G_r(x, y) &\leq G_T, \text{ so } w_i (i = 1, 2) \text{ does not change} \\
G_s(x, y) - G_r(x, y) &> G_T \land (w_1 \geq w_2), \text{ so } w_1 = 1, w_2 = 0 \\
G_s(x, y) - G_r(x, y) &> G_T \land (w_1 < w_2), \text{ so } w_1 = 0, w_2 = 1
\end{align*}
\]  

(12)
Using this image mosaic strategy not only takes into account the differences in the overlap between the registered image and the existing mosaic result, but also particularly conforms to the human visual characteristics. Achieve a good smooth transition effect. 4.2 Registration experiment to achieve image matching In this part of the electronic image sequence mosaic method proposed in this article, the C++ programming language is used to ensure the effectiveness of the method, and the related algorithms in this article are successfully completed. At the same time, real electronic images are used to verify the algorithm. From the SIFT feature points obtained in this article 3 (as shown in Figures 1 and 2), the stitched image is obtained through the registration algorithm, as shown in Figure 3.

![Figure 3. Registration image](image)

5. Conclusions
Based on the obtained image feature points, this paper analyzes the image stitching and registration method based on Scale Invariant Feature Transform (SIFT) features, and makes improvements to achieve accurate and robust image registration. Further realized the color fusion method based on visual characteristics to obtain a smooth mosaic effect. Finally, the experiment completed the mosaic image, which verified the effectiveness of the method.

References
[1] Zhou Guoku. Remote sensing image registration technology based on wavelet extraction of edge feature points [D]. Xi'an: Northwestern Polytechnical University, 2002: 30-35 [20] Steve Seitz. Computer Vision: Edge Detection[R]. University of Washington 2002
[2] Lowe,D. Distinctive Image Features from Scale-Invariant Keypoints [J]. International Journal of Computer Vision 2004
[3] Vishyjit S, Nalwa. A Guided Tour of Computer Vision. Addison-Wesley, 2013
[4] Conover,W.J, Practical Non Parametric Statistics[j]. Wiley series in probability and mathematics statistics, 1980(2)
[5] Brown. L G A Survey of Image Registration Techniques[J]. Computing Surveys, 2012 44(4):325-376
[6] Naya, S. Catadioptric Omni Directional Camera[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017, 482-488
[7] Dong Benzhi, Long Jianyong, Jing Weipeng. PKD-tree: An effective SIFT image feature matching method [J]. Computer Engineering and Applications, 2017, 53(16): 182-186