

A Fast Image Matching Algorithm Based on Yolov3

LIU Rui, LENG Xuefei*, LIU Yang

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P. R. China

(Received 7 August 2020; revised 28 January 2021; accepted 13 April 2021)

Abstract: In view of the fact that the traditional Hausdorff image matching algorithm is very sensitive to the image size as well as the unsatisfactory real-time performance in practical applications, an image matching algorithm is proposed based on the combination of Yolov3. Firstly, the features of the reference image are selected for pre-training, and then the training results are used to extract the features of the real images before the coordinates of the center points of the feature area are used to complete the coarse matching. Finally, the Hausdorff algorithm is used to complete the fine image matching. Experiments show that the proposed algorithm significantly improves the speed and accuracy of image matching. Also, it is robust to rotation changes.

Key words: Yolov3; image matching; Hausdorff; two-stage matching

CLC number: TN967.2 **Document code:** A **Article ID:** 1005-1120(2021)05-0807-09

0 Introduction

Image matching navigation technology is an autonomous navigation method based on image matching technology for precise navigation of carriers. It has broad application prospects in the fields of precision attack, unmanned aerial vehicle (UAV) navigation, deep space exploration, etc. However, autonomous navigation system has special demands for real-time accuracy and robustness of image matching algorithm^[1].

Image matching algorithm includes grayscale matching and feature matching, among which feature matching algorithm is more suitable for autonomous navigation system. In terms of feature-based image matching algorithms, the domestic and foreign scholars have proposed many algorithms, such as the image matching algorithm based on the scale-invariant feature transform (SIFT) proposed by Lowe in 1999^[2]. SIFT can solve the scale and rotation problems, but its use of feature operators seriously slow down the speed of operation. Bay et al.^[3] proposed the speeded up robust feature (SURF) algorithm in 2006. The feature descriptor constructed

by SURF is different from SIFT, but its overall idea is similar to that of SIFT. The computing speed of SURF is about three times that of SIFT. In 2011, Ruble et al.^[4] proposed a fast feature point extraction and description algorithm called oriented fast and rotated brief (ORB), which greatly improves the running speed of the algorithm. In 2017, Zahra et al.^[5] proposed a method to eliminate redundant key points in view of the problem that feature points extracted by SIFT are too close to each other, so as to improve the running factors of the algorithm. The above algorithms are sensitive to the image size so that the matching speed is decreased. The navigation task requires relatively high real-time matching algorithm. The huge amount of data computation will seriously slow down the speed of the process of image matching. Even the fast ORB algorithm can hardly meet the navigation requirements when it is at a huge magnitude of computation^[6].

In 2012, image recognition was realized for the first time through the convolutional neural network (CNN), which promoted the development of deep learning in the field of images. Since then, various

*Corresponding author, E-mail address: lengxuefei@nuaa.edu.cn.

How to cite this article: LIU Rui, LENG Xuefei, LIU Yang. A fast image matching algorithm based on Yolov3[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2021, 38(5): 807-815.

<http://dx.doi.org/10.16356/j.1005-1120.2021.05.009>

CNN frameworks were extensively studied and applied^[7]. Compared with the artificial feature operator, the feature extraction and target detection algorithm based on deep learning can extract the information of the image more intuitively and steadily, and reduce the number of features^[8]. In 2014, Ross et al.^[9] proposed the region-CNN (R-CNN) network. By extracting the candidate regions and classifying the corresponding regions, it played a good role in image recognition. Therefore, in recent years, a variety of algorithms based on R-CNN have also been proposed, such as fast R-CNN, faster R-CNN, and mask R-CNN^[10-11]. Yolo (You only look once), a target detection system based on a single neural network, was proposed by Joseph et al. in 2016^[12], which further improved the accuracy and speed of detection. After that, the author proposed Yolo9000 and Yolov3^[13-14]. The detection speed of these algorithms is improved. MatchNet proposed by Han et al.^[15] and DeepCompare proposed by Zagoruyko et al.^[16] judge whether the images match or not, through the similarity prediction of images based on deep learning. Fan et al.^[17] proposed a spatial-scale double-channel deep CNN method and Tian et al.^[18] proposed L2-NET, both of which matched the tags of the feature areas through deep learning. However, the definition accuracy of these algorithms does not meet the requirement for pixel level accuracy. It is difficult to satisfy the accuracy of navigation.

In view of the special requirements of navigation system for the accuracy and real-time capability of image matching algorithms, this paper divides image matching into two steps: Coarse matching and fine matching based on the idea of two-stage matching. Firstly, on the strength of Yolov3, we detect the feature area of the target recognition model, which can effectively avoid the operation delay caused by the over image size and improve the real-time performance of the image matching algorithm. At the same time, it can effectively remove the influence of image background on image matching and improve the accuracy of matching. Then, we apply the partial Hausdorff distance (PHD) image matching algorithm for matching position correction,

which can avoid the error of the coordinates of the feature area determined by deep learning and improve the accuracy of the image matching algorithm.

1 Coarse Matching Based on Deep Learning

1.1 Yolov3 network structure

Yolov3 extracts feature information from input image by feature extraction network, and gets the feature information of different scales of input image. The first part of Yolov3 network is Darknet-53. The network is from 0 layer to the 74th layer, of which 53 layers are convolution layers. The network extracts the input image features by convolution kernel of 3×3 and 1×1 , and the rest is residual network, which makes the model more easily optimized and more easily access to depth information^[19]. The Yolov3 multi-scale detection network is shown in Fig.1.

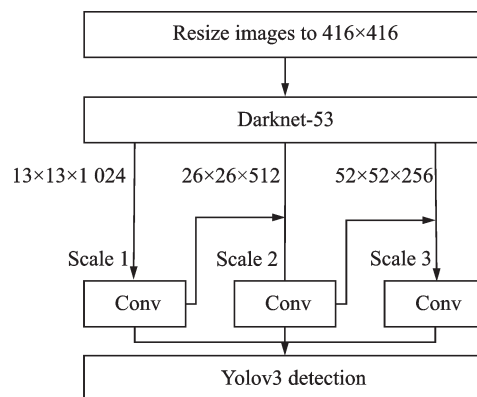


Fig.1 Yolov3 multi-scale detection

From the 75th layer to the 105th layer, the feature information interaction network of Yolov3 network is equivalent to the fully connection layer of the traditional neural network. The network is divided into three scales. The minimum input scale Yolo layer of Yolov3 is 13×13 feature image, containing 1 024 channels altogether. The input of the meso-scale Yolo layer is 26×26 feature image, containing 512 channels altogether. The maximum input scale Yolo layer is 52×52 feature image, containing 256 channels altogether. The local feature interaction is achieved by means of convolution within each scale. Convolution operation does not change the scale of the feature image, while it reduces the

number of feature image channels to n . The size of n is determined by the classification type. Finally, the output of $13 \times 13 \times n$ feature images is carried out to conduct regression and classification.

1.2 Prediction of bounding box

Yolov3 uses K -means clustering to get the size of the prior frame and sets up three prior frames for each of the down sampling scales. There are nine kinds of prior frames which are clustered. In this algorithm, the priori frames provided by COCO dataset are (10×13) , (16×30) , (33×23) , (30×61) , (62×45) , (59×119) , (116×90) , (156×198) , (373×326) , respectively. The minimum 13×13 characteristic map is applied with the maximum prior frames (116×90) , (156×198) , (373×326) , which are suitable for detecting larger objects. The median 26×26 characteristic map is applied with the median prior frames (30×61) , (62×45) , (59×119) , which are suitable for detecting medium-sized objects and detecting smaller objects. The maximum 52×52 characteristic map is applied with the minimum prior frames (10×13) , (16×30) , (33×23) , which are suitable for detecting smaller objects.

The algorithm uses the intersection over union (IOU) to measure the accuracy of detecting corresponding objects in a specific data set. When the statistical IOU is below 0.5, the prediction of the algorithm is considered to be inaccurate. Bounding-Box regression is used to fine tune the window to make the inaccurate feature box closer to the correct mark. So that the location will be more accurate. The coordinate information of the characteristic region is obtained. The specific mapping relations are as follows

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

where t_x, t_y, t_w, t_h are the actual output targets of the Yolov3 model as the input of the window mapping transformation; $\sigma(t)$ is Sigmoid function, as a criterion for transformation; b_x, b_y, b_w, b_h are coordinates and sizes of the center of bounding box; c_x, c_y are the corresponding coordinates of feature image; p_w, p_h are the corresponding coordinates of prior frame.

1.3 Dataset and experiments

Satellite images and the features are selected. By rotating any angle and adding random noise, including the pepper and salt noise and the Gaussian noise, the data are further enhanced by random scale transformation between 0.8 and 1.4, and the dataset is expanded to 2 000 pieces, as shown in Fig.2. Part of dataset is shown in Fig.3.

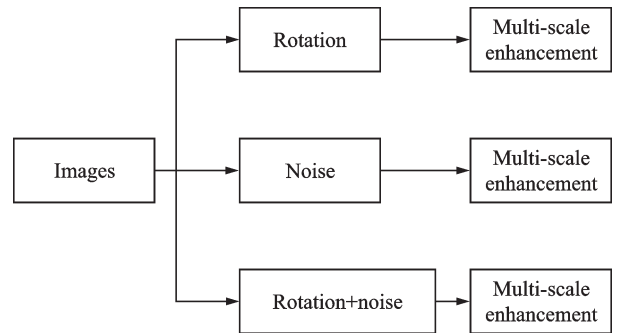


Fig.2 Schematic diagram of data enhancement

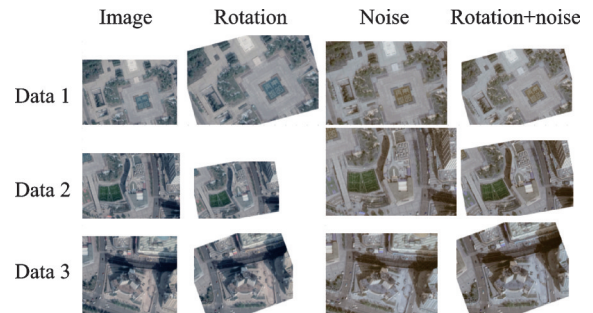


Fig.3 Part of training set samples

After completing the dataset, the Yolov3 network is used to train and iterate 2 000 times. This experiment runs on the computer with Ryzen 3600 CPU, 16 GB RAM and RTX2060 GPU, 6 GB RAM; the operation system is LinuxUbuntu16.04; Python2.7 is compiled; the main algorithm frame is Darknet and OpenCV2. After the training is completed, the weight file is tested and verified. The accuracy rate of the test set is 98.36%. The detection time of the single picture is 0.094 s. Experimental results are shown in Fig.4.

In this paper, other datasets are also established and the results of the detection are verified. Some of the results are shown in Fig.5.

The prediction label will be used as the basis for judging coarse matches. If the labels are consis-

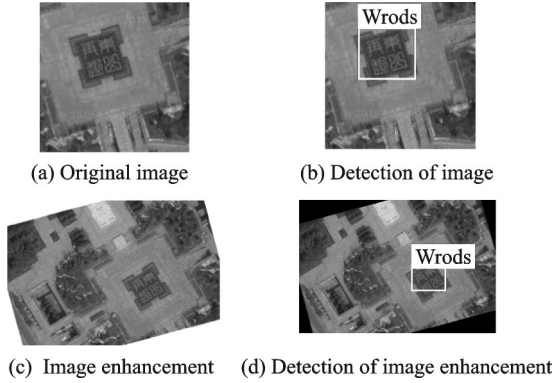


Fig.4 Results of Yolov3 image detection

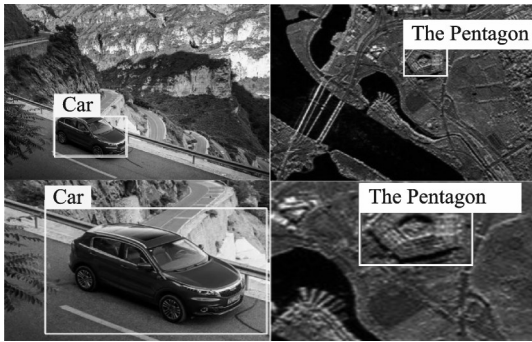


Fig.5 Results of Yolov3 image detection on other datasets

tent, the feature areas extracted from the two images are considered to be matched successfully. Then, the next fine matching is carried out.

2 Fine Matching Based on PHD Image Matching Algorithm

Hausdorff distance is a way to describe the least similarity between two images. PHD image matching algorithm can avoid the influence of noise and false detection points. PHD image matching algorithm is introduced to carry out two-stage fine matching.

PHD between the given image A and the image B is defined as^[20]

$$\begin{cases} H_{\text{PHD}}(A, B) = \max(h_k(A, B), h_l(A, B)) \\ h_k(A, B) = k_{a \in A}^{\text{th}} d(a, B) \quad k = f \times N_A, 0 < f < 1 \\ h_l(B, A) = k_{a \in A}^{\text{th}} d(b, A) \quad l = g \times N_B, 0 < g < 1 \end{cases} \quad (5)$$

where f and g are forward and backward coefficients respectively; N_A, N_B represent the total numbers of points in images A and B , respectively; $d(a, B) = \min_{b \in B} \|a - b\|$ represents the minimum distance

from point b in set B to point a ; $d(b, A) = \min_{a \in A} \|a - b\|$ represents the minimum distance from point a in set A to point b . The Hausdorff distance is the least similarity between the two images. By traversing all the feature points, the characteristic point corresponding to the smallest Hausdorff distance is found to be the best matching point.

PHD image matching algorithm is of high matching rate and simple computation logic. But from the algorithm flow, it can be found that the computation of the algorithm is determined by the number of feature points of the image, so the algorithm is very sensitive to the complexity and size of the image. Therefore, the Yolov3 algorithm can be used to complement each other and to accelerate the Hausdorff algorithm. At the same time, the Hausdorff algorithm is used to perform the two-stage fine image matching.

3 Two-Stage Image Matching Algorithm Flow

In this paper, a two-stage image matching algorithm combining Yolov3 algorithm and Hausdorff image matching algorithm is proposed. Firstly, the coarse matching is carried out by the Yolov3 algorithm, and then the two-stage fine matching is made by using the PHD image matching algorithm. Meanwhile, due to the large image size, the PHD image matching algorithm does not meet the requirement for real-time in image matching navigation. The Yolov3 algorithm can improve the real-time performance of the algorithm and satisfy the requirement of image matching navigation.

Although there are many convolutional layers, the convolutional layers of many channels have no inheritance. In addition, the estimation of feature area prediction becomes simpler. On one hand, each feature region only matches one prior box, which reduces the complexity of the algorithm. On the other hand, the time complexity of the Hausdorff image matching algorithm is $O(n, m)$, which means its complexity is related to the feature points of pre-stored and measured graphs. The coarse matching

can greatly reduce the number of feature points and greatly reduce the complexity of the algorithm. The detailed flow chart of the algorithm is shown in Fig.6.

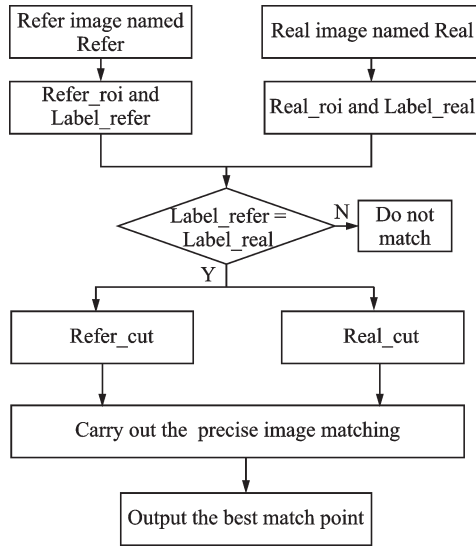


Fig.6 Flow chart of two-stage matching algorithm

The specific process is as follows:

(1) Read in the refer image named Refer and the real image named Real. Use the Yolov3 algorithm to extract the feature image from the refer image and the real image to achieve the feature regions named Refer_roi and Real_roi, respectively and corresponding region labels named Label_refer and Label_real, respectively.

(2) If Label_refer and Label_real are consistent, the next step will be fine matching. Otherwise, the two images do not match.

(3) Label_refer and Label_real correspond to Refer_roi and Real_roi, respectively. By using the center point coordinates as the center, intercept Refer_cut and Real_cut from Refer_roi and Real_roi, respectively.

(4) Use Refer_cut and Real_cut as the input of PHD image matching algorithm to carry out the fine image matching.

(5) Output the best matching point between Refer_cut and Real_cut.

(6) Use the best matching point output in the 5th step to the point corresponding to the original images, and output the best match point.

4 Experiment

In order to verify the effectiveness of the algorithm, a matching experiment is carried out. This paper validates the algorithm for ordinary optical images, synthetic aperture radar (SAR) images and satellite images. The experiment runs on the computer with Ryzen 3600 CPU 16 GB RAM and RTX2060 GPU 6 GB RAM, the operation system is LinuxUbuntu16.04, Python2.7 is compiled, and the compilation environment is pycharm2018.1. The main algorithm framework is Darknet and opencv2.

4.1 Coarse matching based on Yolov3

Select the Washington SAR image and choose the Pentagon as the feature. Besides, we randomly intercept 500 images and expand the dataset by rotating arbitrary angle and adding Gauss noise and salt and pepper noise to 2 000. Part of dataset is shown in Fig.7.

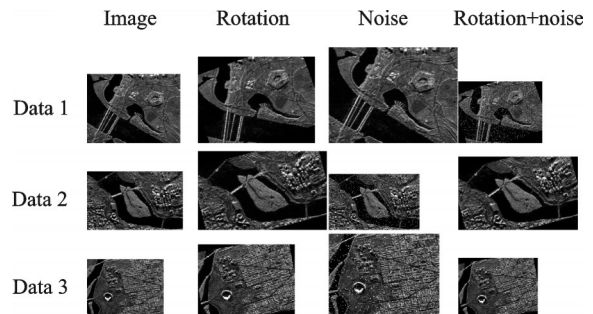


Fig.7 Schematic diagram of image enhancement

After completing the dataset, the Yolov3 network is used for training. The number of iterations is 2 000 times. After the training is completed, the weight files are tested and verified. The accuracy rate of the test set is 98.36%, and the detection time of the single picture is 0.094 s.

We select the refer image named Refer with the size of 329 pixel \times 214 pixel and the real image named Real with the size of 130 pixel \times 100 pixel, as shown in Fig.8. Using the trained model for feature detection, we can predict the bounding boxes as the feature area named Refer_roi of the refer image and Real_roi of the real image. These images are shown in Fig.9. The labels of the two feature ar-

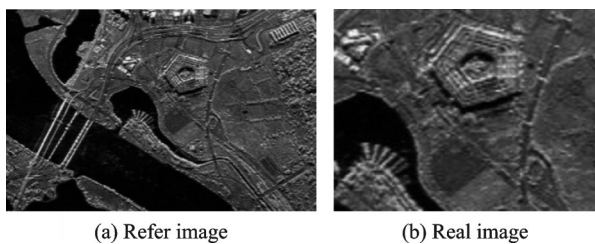


Fig.8 Refer image and real image

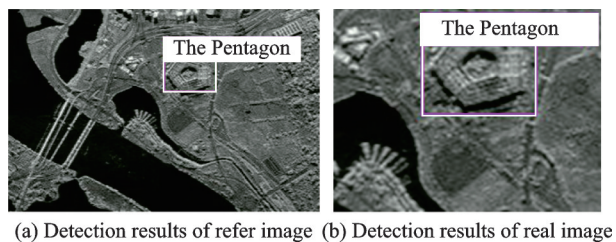


Fig.9 Detection results of refer and real images

eas are the same, so the center points of the two feature areas are the coarse matching results for fine matching.

4.2 Fine matching based on PHD image matching algorithm

Beside taking the center point coordinates of bounding box as the center, intercept the two-stage feature images from Refer_roi with size of $40 \text{ pixel} \times 40 \text{ pixel}$ and Real_roi Refer_roi with size of $30 \text{ pixel} \times 30 \text{ pixel}$. The two images, called Refer_cut and Real_cut, are shown in Fig.10. Using the intercepted images as the input of PHD image matching algorithm, the experiment shows that the exact matching point is (2, 4) and the matching time is 0.275 s. The matching result of Refer_cut and Real_cut images is shown in Fig.11.

The final matching point of the algorithm is (120, 40) and the matching time is 0.476 s. Result is shown in Fig.12.

If only Yolov3 is used as the matching algorithm, the matching point is (123, 44) and the cor-

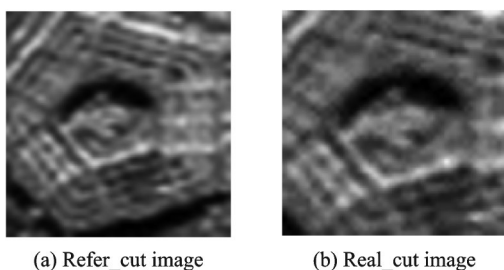


Fig.10 Refer_cut and Real_cut images



Fig.11 Matching result of Refer_cut and Real_cut images



Fig.12 Final matching result of Refer and Real images

rect matching point is (120, 40) with a matching error of 5. From the experimental results, it can be seen that if only Yolov3 is used for image matching, the matching error does not meet the requirement of matching accuracy. If only the Huasdorff matching algorithm is used, the matching error is 0 and the matching time is 12 s, which satisfies the accuracy requirement but does not satisfy the real-time performance of the algorithm. When the two algorithms are combined, the matching error is 0 and the matching time is 0.476 s, which means that the real-time and accuracy requirements are satisfied at the same time.

Similarly, the above algorithm is applied to optical images and satellite images. The total matching time of optical images is 0.476 s, and the final matching point is (99, 357). The matching result is shown in Fig.13. The total matching time of the satellite image is 0.391 s, and the final matching point is (120, 40), as shown in Fig.14.

From the matching results, we can find that the algorithm has high matching accuracy in all kinds of images, and the total time of matching is less than 1 s, which meets the real-time requirement. The statistical results of the two-stage matching algorithm are shown in Table 1. From Table 1,

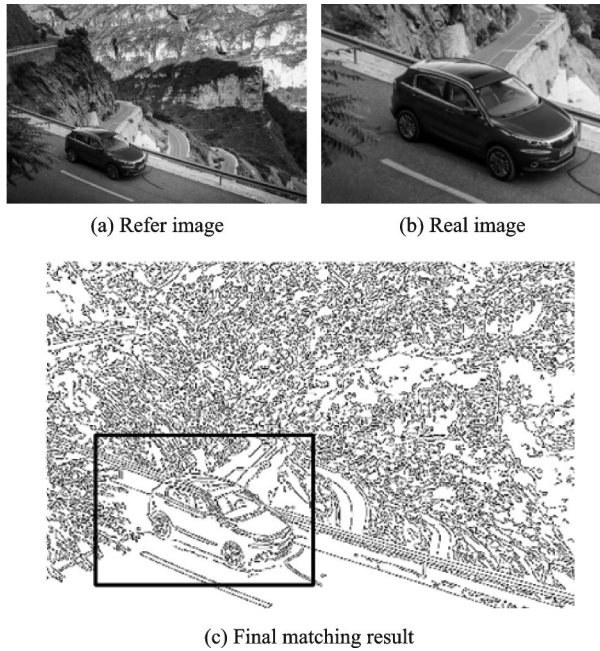


Fig.13 Matching results of optical images

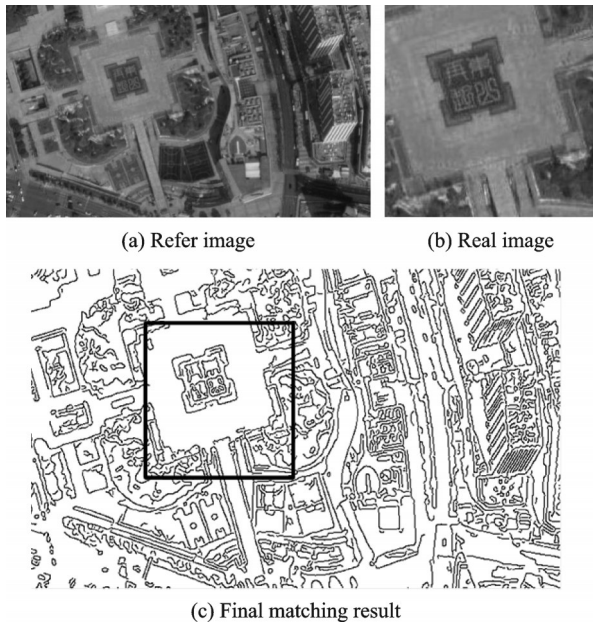


Fig.14 Matching results of satellite images

we can also find that the matching error of the standard algorithm in a certain angle rotation is within 1 pixel, which meets the matching requirements of navigation. Moreover, it is proved that the algorithm is robust.

The proposed algorithm is compared with image matching algorithms based on SURF, ORB and wavelet transform (WT). In the experiment, the matching time of WT based two-level image matching algorithm is about 10 s, which is significantly

Table 1 Statistical results of the two-stage matching algorithm

Image	Size/ (pixel \times pixel)	Angle/ ($^{\circ}$)	Point	Error	Time/s
SAR image	$1\ 080 \times 720$	0	(15,4)	0.000	0.476
	$1\ 080 \times 720$	5	(15,4)	0.000	0.482
	$1\ 080 \times 720$	10	(17,5)	2.236	0.489
Optical image	329×214	0	(2,4)	0.000	0.334
	329×214	5	(2,4)	0.000	0.345
	329×214	10	(4,5)	1.414	0.339
Satellite image	702×410	0	(4,5)	0.000	0.391
	702×410	5	(4,5)	0.000	0.402
	702×410	10	(5,6)	1.414	0.399

higher than other algorithms. The matching time of the remaining three types of algorithms is shown in Fig.15.

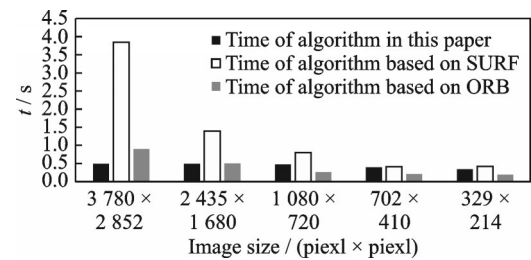


Fig.15 Statistical chart of matching error under rotation

From the contrast results, it can be found that the real-time performance of algorithm in this paper and the image matching algorithm based on ORB are obviously better than the image matching algorithm based on SURF. At the same time, the matching time of the image matching algorithms based on ORB and SURF increases obviously with the increase of image size. However, with the increase of image size, the matching time of the proposed algorithm is less than 0.5 s, which is better than the image matching algorithm based on ORB.

Considering that the ORB algorithm is similar to the proposed algorithm in terms of time performance, only the performance of these two algorithms is compared. Fig.16 shows the matching results of the $1\ 080\ \text{pixel} \times 720\ \text{pixel}$ image using ORB. From the result, it can be found that only one correct matching point exists among the seven successful matching points. And image matching accuracy comparison is shown in Fig.17. From Fig.17, it can be found that the image matching algorithm

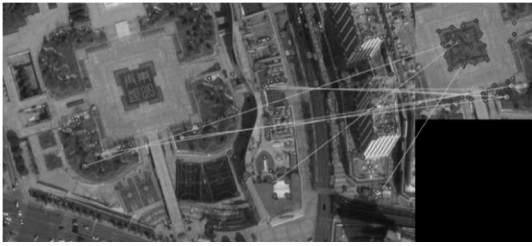


Fig.16 Matching results based on ORB

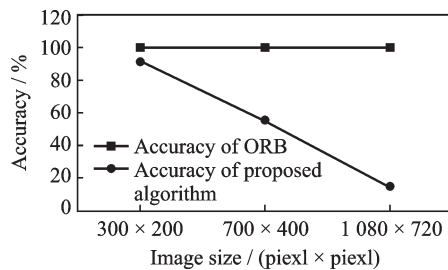


Fig.17 Accuracy comparison of image matching algorithms

based on ORB has obvious matching error in the actual image application in some cases. In the case of similar matching time, the proposed algorithm has the advantage of high matching accuracy.

5 Conclusions

This paper first introduces the structures of YOLOv3 CNN and PHD image matching algorithms. Aiming at solving the problems that there is deviation in the bounding box predicted by YOLOv3 CNN, which does not meet the accuracy requirements, and PHD image matching algorithm is sensitive to image size, which does not meet real-time requirements, we make a targeted improvement by the combination of the two algorithms. Experimental results show that the algorithm is effective, accurate and real-time. At the same time, it can resist a certain degree of rotation. However, in the actual situation, how to match the incomplete features and how to distinguish two similar features still need to be solved.

References

- [1] JIA D, ZHU N D, YANG N H, et al. A survey of image matching methods[J]. Chinese Journal of Image and Graphics, 2019, 24(5): 677-699.
- [2] LOWE D G. Object recognition from local scale-invariant features[C]//Proceedings of the Seventh IEEE International Conference on Computer Vision, Greece: IEEE, 1999, 2: 1150-1157.
- [3] BAY H, TUYTELAARS T, VAN GOOL L. SURF: Speeded up robust features[C]// Proceedings of European Conference on Computer Vision. Berlin: [s.n.], 2006, 3951: 404-417.
- [4] RUBLE E, RABAU D V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]//Proceedings of International Conference on Computer Vision. Barcelona: [s.n.], 2011: 2564-2571.
- [5] ZAHRA H N, MEHDI N. RKEM: Redundant key-point elimination method in image registration[J]. IET Image Processing, 2017, 11(5): 273-284.
- [6] YU Qiuze. Synthetic aperture radar (SAR) image matching navigation technology[D]. Wuhan: Huazhong University of Science and Technology, 2004. (in Chinese)
- [7] ALEX K, I SUTSKEVER, G HINTON. ImageNet classification with deep convolutional neural networks[J]. Association for Computing Machinery, 2017, 60(6): 84-90.
- [8] WANG E S, SONG Y S, Xu S, et al. ADS-B anomaly data detection model based on deep learning and difference of Gaussian approach[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2020, 37(4): 550-561.
- [9] ROSS G, JEFF D, TREVOR D, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2014: 580-587.
- [10] REN S, HE K, ROSS G, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [11] HE K, GRORGIA G, PIOTR D, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2961-2969.
- [12] JOSEPH R, SANTOSH D, ROSS G, et al. You only look once: United, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2016: 779-788.
- [13] JOSEPH R, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08) [2020-02-20]. <https://arxiv.org/abs/1804.02767>.
- [14] JOSEPH R, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Confer-

- ence on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 7263-7271.
- [15] HAN X, LEUNG T, JIA Y Q, et al. MatchNet: Unifying feature and metric learning for patch-based matching[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 3279-3286.
- [16] ZAGORUYKO S, KOMODARKIS N. Learning to compare image patches via convolutional neural networks[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 4353-4361.
- [17] FAN D Z, DONG Y, ZHANG Y S. Satellite image matching method based on deep convolution neural network[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(6): 844-853.
- [18] TIAN Y R, FAN B, WU F. L2-Net: Deep learning of discriminative patch descriptor in Euclidean space[C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern. Honolulu, USA: IEEE, 2017: 661-669.
- [19] TAN Jun. An improved YOLOv3 target recognition algorithm[D]. Wuhan: Huazhong University of Science and Technology, 2018. (in Chinese)
- [20] LENG Xuefei. Key technologies of image matching aided navigation system based on image features[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2007. (in Chinese)

Acknowledgement This work was supported by the Foundation of Graduate Innovation Center in Nanjing University of Aeronautics and Astronautics (No.kfjj20191506).

Authors Mr. LIU Rui received the B.S. degree in Electrical Engineering and Automation from Southwest Jiaotong University. Currently he is studying for the M.S. degree in College of Astronautics, Nanjing University of Aeronautics and Astronautics. His research direction is focused on scene matching navigation.

Dr. LENG Xuefei received the B.S. degree and Ph.D. degree in Measurement and Control Technology and Instrumentation Program from Nanjing University of Aeronautics and Astronautics in 2000 and 2007, respectively. She joined in Nanjing University of Aeronautics and Astronautics in 2007, where she is an associate professor of College of Astronautics. Her research is focused on integrated navigation system and image processing fields.

Author contributions Mr. LIU Rui designed the study, compiled the models, conducted the analysis, interpreted the results and wrote the manuscript. Dr. LENG Xuefei gave guidance on direction guidance and experimental protocols. Mr. LIU Yang contributed to the discussion and background of the study. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: ZHANG Huangqun)

基于 Yolov3 神经网络的快速图像精匹配算法

刘 瑞, 冷雪飞, 刘 杨

(南京航空航天大学航天学院, 南京 211106, 中国)

摘要:针对传统的 Hausdorff 图像匹配算法对于图像尺寸十分敏感、在实际应用中实时性不强的问题,提出一种结合 Yolov3 神经网络的图像匹配算法。首先,选取参考图像中的特征进行预训练;然后,利用训练结果对实测图像进行特征提取,利用特征区域的中心点坐标完成粗匹配;最后,对参考图与实测图的特征区域进行截取,利用 Hausdorff 算法完成图像的精匹配。实验结果表明,本文提出的算法显著提高了运算速度及匹配准确率,并且对旋转变换具有一定的鲁棒性。

关键词: Yolov3; 图像匹配; Hausdorff; 二级匹配