

Artificial Intelligence Registration of Image Series Based on Multiple Features

Zhixin Li, Degang Kong*, Yongchun Zheng

School of Computer Technology and Engineering, Changchun Institute of Technology, Changchun 130012, China

Corresponding Author Email: rj_lzx@ccit.edu.cn



<https://doi.org/10.18280/ts.390122>

ABSTRACT

Received: 2 November 2021

Accepted: 10 January 2022

Keywords:

multi-source image series, image registration, image feature extraction

Multi-source image series vary in quality. To fuse the feature information of multi-source image series, it is necessary to deeply explore the relevant registration and fusion techniques. The existing techniques of image registration and fusion lack a unified multi-feature-based algorithm framework, and fail to achieve real-time accurate registration. To solve these problems, this paper probes into the artificial intelligence (AI) registration of image series based on multiple features. Firstly, the Harris corner detector was selected to extract the corners of multi-source image series, before explaining and improving the flow of the algorithm. In addition, the deep convolutional neural network (DCNN) VGG16 was improved to extract the features of multi-source image series. Finally, the spatial transformation network was adopted to pre-register the image series, and the image series was deformed and restored based on the region-constrained moving least squares. The proposed registration algorithm was proved effective through experiments.

1. INTRODUCTION

Image registration, which is crucial to image processing, can combine the image information from multiple sources, and help people understand the current image or scene more clearly and comprehensively, thereby meeting various practical needs [1-9].

The same shooting equipment can generate multiple frames of images. The multi-source image series produced by different shooting equipment from different perspective at different time vary in quality [10-17]. Through comprehensive analysis of multi-source image series, it is possible to synthesize the image feature information of interest for future research [18-20]. Therefore, it is of great theoretical and practical significance to study the registration and fusion techniques of multi-source image series.

In the field of medical image processing, Dolly et al. [21] stabilized the medical images acquired from different modes to eliminate jitter artifacts, and further fused the treated images, enabling doctors to visualize the composite features of computed tomography (CT) and magnetic resonance imaging (MRI). Each pair of images were aligned to ensure parameter registration, which in turn supports effective mixing and superposition and guarantees diagnosis correctness. Mingjun et al. [22] proposed a fast, memory-efficient registration algorithm for image series-based computer visual applications: Firstly, a series of feature points were generated by a new scheme unrelated to image contents; Then, reliable pairs of feature points were obtained through forward and backward tracking. Chen et al. [23] improved the feature-based series comparison approach, and combined it with the feature-based image registration method. Experiment results show that the combined strategy increased the alignment accuracy of the X-axis and Y-axis. The effectiveness of the strategy was also demonstrated by applying the registration results to spatiotemporal super-resolution reconstruction.

In recent years, multi-mode cameras, consisting of red-green-blue (RGB) sensors and infrared (IR) sensors, are increasingly adopted in monitoring and robot applications. The traditional computer vision method, which maps image pairs using pixel intensity or image features, is not suitable for RGB/IR image pairs. With the aid of variational optimization, Kirby and Whitaker [24] mapped the optical flow field calculated from different wavelength images, making up for the lack of common features in RGB/IR image pairs. Their approach was tested on synthetic optical flow fields, and the actual image series created by multi-mode binocular stereo RGB/IR cameras. Fan et al. [25] proposed a rapid and effective registration algorithm for drone image series. The algorithm covers three main steps, namely, feature extraction, feature point tracking, and matrix estimation. A series of experiments were conducted on different image series. The experimental results reveal the good image registration effect of the algorithm: the mean registration time was 0.3s.

After many years of development, image series registration has yielded many research results. But lots of problems are yet to be addressed: wrong feature matching abounds, due to the presence of every edge information of the original image; no unified algorithm framework is available based on multiple features. Meanwhile, the constant updates of image shooting equipment, as well as the development of image processing techniques, raise more and more strict requirements on the real-time performance and automation of image series registration. Therefore, this paper probes into the artificial intelligence (AI) registration of image series based on multiple features. Section 2 selects the Harris corner detector to extract the corners of multi-source image series, and details the original and improved schemes of the algorithm. Section 3 extracts the features of multi-source image series with the deep convolutional neural network (DCNN) VGG16, pre-register the image series by the spatial transformation network, and deforms and restores the image series based on the region-

constrained moving least squares. Experimental results confirm the effectiveness of the proposed registration algorithm [26, 27].

2. FEATURE POINT EXTRACTION

When it comes to feature-based image registration, one of the difficulties is how to extract features like feature points, edge features, and regional features. The effectiveness of image features is critical to the accurate registration of multi-source image series, and exerts a huge effect on subsequent tasks of image processing.

Corners play a decisive role in obtaining the target contours of multi-source image series. The classic detection methods include Moravec corner detector, SUSAN (smallest univalue segment assimilating nucleus) corner detector, MIC (minimum change of intensity) corner detector, and Harris corner detector. Among them, Harris corner detector was selected to extract the corners of multi-source image series, by virtue of its excellent performance on the noise map.

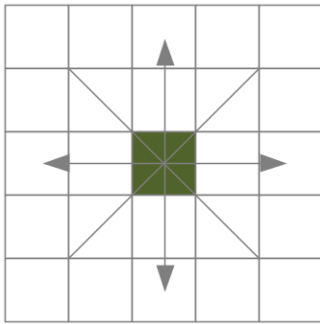


Figure 1. Four directions of the target pixel

Harris corner detector is an improvement of Moravec corner detector, which can be implemented in two steps. The first step is to compute the value of interest INT for each pixel. Figure 1 shows the four directions of the target pixel. In the $\theta \times \theta$ window centering on the target pixel (d, s), the quadratic sum of the grayscale difference between the adjacent pixels in the four directions of the target pixel can be calculated by:

$$\left\{ \begin{array}{l} INT_{d,s} = \min \{u_1, u_2, u_3, u_4\} \\ U_1 = \sum_{i=-l}^{l-1} (h_{d+i,s} - h_{d+i+1,s})^2 \\ U_2 = \sum_{i=-l}^{l-1} (h_{d+i,s-i} - h_{d+i+1,s+1+i})^2 \\ U_3 = \sum_{i=-l}^{l-1} (h_{d,s+i} - h_{d,s+1+i})^2 \\ U_4 = \sum_{i=-l}^{l-1} (h_{d+i,s-i} - h_{d+i+1,s-1-i})^2 \end{array} \right. \quad (1)$$

where, $l = \text{int}(\theta/2)$. The minimum among the quadratic sums in the four directions is selected as the INT of the target pixel (d, s). Next is to set the threshold, and define the candidate corner as a pixel with an INT greater than the preset threshold.

The defects of Moravec corner detector are solved to obtain the Harris corner detector. Firstly, the grayscale variation of the image window is expanded based on the first-order derivative, so that it can fully characterize the grayscale variation in each direction. Besides, Gaussian window weighting is performed to minimize the interference of image noise. The updated grayscale variation of the image window can be calculated by:

$$\begin{aligned} R_{a,b} &= \sum_{v,u} \theta_{v,u} [SU_{a+v,b+u} - SU_{v,u}]^2 \\ &= \sum_{v,u} \theta_{v,u} [aA + bB + p(a^2 + b^2)]^2 \end{aligned} \quad (2)$$

The first-order derivative can be approximately expressed as $A = SU \otimes (-1, 0, 1) = \partial SU / \partial a$, $B = SU \otimes (-1, 0, 1) = \partial SU / \partial b$, where $\theta_{v,u} = r^{-(v^2+u^2)/(2\sigma^2)}$ is the standard Gaussian function. Then, formula (2) can be simplified as:

$$R_{a,b} = Xa^2 + 2Zab + Yb^2 \quad (3)$$

Let \otimes denote convolutional operation. We have:

$$\left\{ \begin{array}{l} X = A^2 \otimes \theta \\ Y = B^2 \otimes \theta \\ Z = (AB) \otimes \theta \end{array} \right. \quad (4)$$

The corner judgement criterion can be redefined as:

$$R_{a,b} = (a, b)N(a, b)^T \quad (5)$$

where, N is a 2×2 symmetric matrix: $N = \begin{bmatrix} X & Z \\ Z & Y \end{bmatrix}$.

It can be found that the change of $R_{a,b}$ is closely correlated with the local autocorrelation function of the image, which carries both direction and texture features. The shape of $R_{a,b}$ at the origin can be characterized by N. In other words, the main curvature of the local autocorrelation function is proportional to the two eigenvalues γ and δ . Therefore, both γ and δ can be used to depict the rotational invariance of image features. Special consideration should be given to the following three cases:

(1) If γ and δ are both small, the function curve is straight, and the brightness of the window area can be regarded as approximately constant;

(2) If one of γ and δ is big and the other is small, then the function curve will be tall in the middle and shallow on the two sides, and can be displayed as an edge;

(3) If γ and δ are both big, the curve would change abruptly from the tall center to the low sides, and exhibit as a corner.

To accurately measure the angle of a pixel, it is necessary to examine and analyze γ and δ . For simplicity of eigenvalue computing, the trace and determinant of N can be given by:

$$\left\{ \begin{array}{l} \text{TRA}(N) = \gamma + \delta = X + Y = SU_A^2 + SU_B^2 \\ \text{DET}(N) = \gamma\delta = XY - Z^2 \\ = SU_A^2 SU_B^2 - (SU_A SU_B)^2 \end{array} \right. \quad (6)$$

Further, the angle response function can be defined. Let l be a constant specific to the Gaussian function and differential template. Then, the angle response function can be defined as:

$$S = DET(N) - l \bullet TRA(N)^2 \quad (7)$$

where, $l = \tau / (1 + \tau^2)$; $1 / \tau < \gamma / \delta < \tau$.

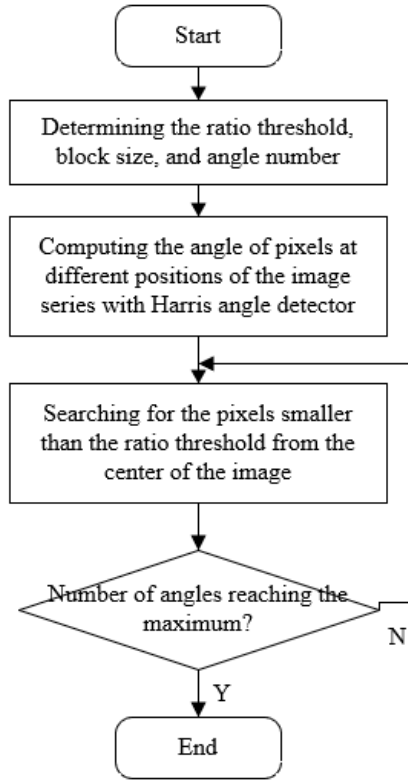


Figure 2. Flow of the improved operator

During the image series matching based on multiple feature points, the most effective way to overcome the limitations of image quality and feature extractors is to screen multiple feature points to be matched. Thus, this paper screens the corners detected by Harris corner detector, using the neighborhood non-maximum suppression (NMS) and global suppression. The goal is to ensure the uniform distribution of angles. The specific flow is shown in Figure 2.

3. AI REGISTRATION

At present, both feature fusion and feature enhancement, capable of making a more reasonable feature description, have exhibited their advantage in image registration. However, these image registration methods are limited by the use of one or two feature points, which contain shallow information only. Neural network-based models provide new solutions to the description of multi-feature images. Therefore, this paper selects the improved DCNN VGG16 to extract the features of multi-source image series. Firstly, transfer learning is adopted to fix the weights and optimize the fully connected layer of the neural network model. Next, the feature points extracted from multiple image series are compiled into a set of control points, which limits the registration range of multi-source image series. In this way, the image registration becomes more precise.

This paper improves the VGG16 by dividing the fully connected layer into two layers. One of them has 4,096 nodes, and the other has 4 nodes. In addition, the original softmax layer is replaced with a softmax classifier with 4 class labels.

3.1 VGG16-based feature extraction

VGG16 extracts features through convolutional operations. Figure 3 explains the process of convolutional operations. Let I_{ow}^k be the kernel of the size o, w on the k -th layer; χ be the bias of each feature mapping; a_{ij}^k be the output of the k -th layer; i, j be the height of the two-dimensional (2D) vector; g be the rectified linear unit (ReLU) activation function. Then, the convolutional operation can be expressed as:

$$a_{ij}^k = g(v_{ij}^k) = g\left(\sum_{o=1}^e \sum_{w=1}^e a_{i+o-1, j+w-1}^{k-1} \times I_{ow}^k + \chi^k\right) \quad (8)$$

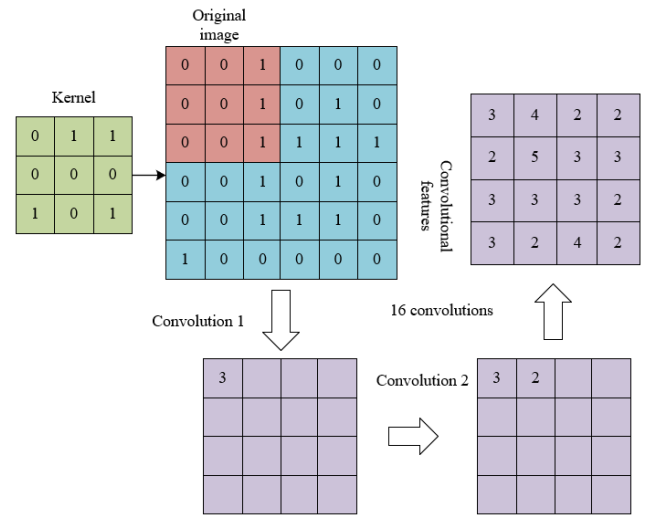


Figure 3. Sketch map of convolutional operations

To give VGG16 the ability of nonlinear mapping, this paper employs ReLU as the activation function, in view of its fast convergence and non-saturability. The nonlinear factor was introduced to the function, aiming to speed up the learning, and reduce the error of backpropagation. The modified ReLU can be expressed as:

$$g_{ReLU}(a) = \begin{cases} 1(a \geq 0) \\ 0(a < 0) \end{cases} \quad (9)$$

When the input is $a \geq 0$, the activation function outputs 1; when the input is $a < 0$, the activation function outputs 0. Besides, max pooling was adopted to reduce the dimensionality of the convolutional outputs, with the aim to reduce the size of image features, simplify the computing process, and retain as much texture information as possible. Let $a_{i,j}$, $a_{i+1,j}$, $a_{i,j+1}$, and $a_{i+1, j+1}$ be the four parameters in the pooling domain; \max be the maximization function. Then, the result of max pooling b_{MP} can be calculated by:

$$b_{MP} = \max(a_{i,j}, a_{i+1,j}, a_{i,j+1}, a_{i+1,j+1}) \quad (10)$$

The last max pooling layer is followed by a fully connected layer and a softmax classifier, which classifies its input into 1

classes. The output of the softmax classifier indicates the probability of an image series sample belonging to a class:

$$E_j = r^{c_j} / \sum_{l=1}^l r^{c_l} \quad (11)$$

During the iterative network training, the network parameters need to be optimized through backpropagation. Let $1\{b_{(i)=j}\}$ be the indicative function. Then, the loss function can be defined as:

$$LOSS_i = - \sum_{j=1}^l 1\{b_{(i)} = j\} \log \frac{PD^{c_j}}{\sum_{k=1}^l PD^{c_k}} \quad (12)$$

The probability for PD^{c_j} to belong to the j -th class can be expressed as:

$$\frac{PD^{c_j}}{\sum_{k=1}^l PD^{c_k}} \quad (13)$$

When the equation within the braces of the indicative function holds, the loss is 1; when that equation does not hold, the loss is 0. Hence, the loss function $LOSS$ only retains one of the terms in the summation operation. At this time, class j is the class of the image series sample:

$$\log \frac{PD^{c_j}}{\sum_{k=1}^l PD^{c_k}} \quad (14)$$

3.2 Pre-registration and registration

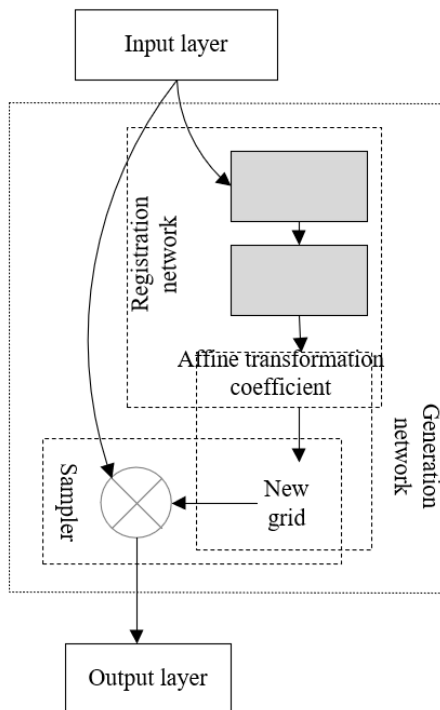


Figure 4. Principle of spatial transformation network

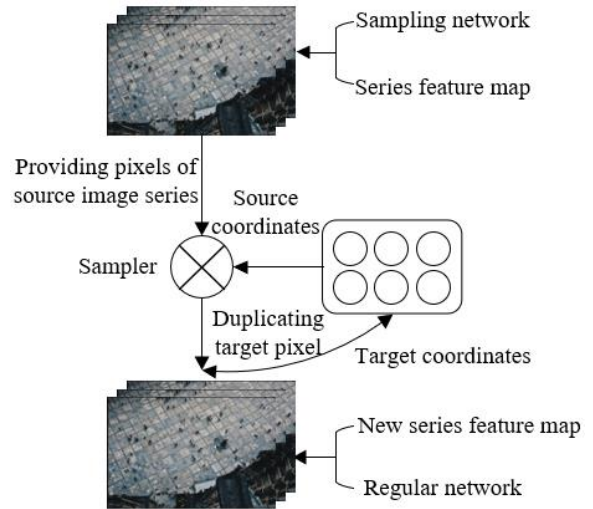


Figure 5. Process of forward propagation

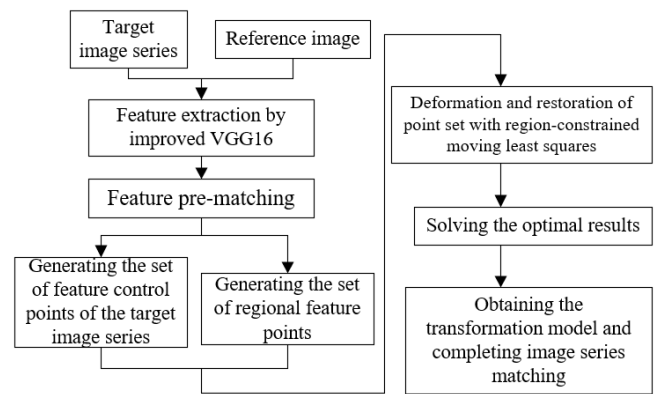


Figure 6. Flow of image series registration

In the multi-feature image series to be registered, some images may undergo deformations like rotation, scaling, and translation, owing to the position change of cameras. Based on templates, the traditional registration methods generally consume a long time to search through the deformed images. This paper pre-registers the image series by the spatial transformation network, aiming to realize the adaptive transformation of the image series and shorten the time of image registration. The principle of the spatial transformation network is explained in Figure 4. It can be seen that the affine transformation coefficient can be generated by the spatial invariance between the target image series and the reference image. The target image series can enter spontaneous geometric change, according to the generated affine transform coefficient, and achieve the effect of pre-registration and initial correction. Figure 5 shows the mechanism of forward propagation.

The region-constrained moving least squares was adopted for deformation and restoration to realize high-quality local registration of multi-source image series, without inducing global deformation of image series. Figure 6 explains the whole flow of image series registration.

Let o be the set of feature control points of the target image series; w be the set of regional feature points; o_i and w_i be the point sets of control points, $o \in \{O_i^*\}$, $w \in \{W_i^*\}$; G be the deformation function; N_{LT} be the linear transformation term; Φ be the translation term. For any pixel u in the image series, a deformation function can be set up $G_u(a) = Na + \Phi$, and simplified for the following optimization problem:

$$\underset{G_u}{\operatorname{argmin}} \sum_i \omega_i |G_u(o_i) - w_i|^2 \quad (15)$$

Let γ be the parameter adjusting the shape change effect. The weight of a point in the control point set o_i can be expressed as:

$$\omega_i = \frac{1}{|o_i - u|^{2\gamma}} \quad (16)$$

The solution to the optimization problem can be expressed as:

$$\begin{cases} N = \left(\sum_i (o_i - o_*)^T \omega_i (o_i - o_*) \right)^{-1} \\ \cdot \sum_i (o_i - o_*)^\Phi \omega_i (w_i - w_*) \\ \Phi = w_* - o_* N \end{cases} \quad (17)$$

Let q_* and o_* be the weighting center. Then, we have:

$$o_* = \frac{\sum_i \omega_i o_i}{\sum_i \omega_i}, w_* = \frac{\sum_i \omega_i w_i}{\sum_i \omega_i} \quad (18)$$

Using the closed solution of N_{LT} , G_u can be updated by:

$$G(u) = (u - o_*) \left(\sum_i o_i'^\Phi \omega_i o_i' \right)^{-1} \sum_j \omega_j o_j'^\Phi w_j' + w_* \quad (19)$$

where, $o_i' = o_i - o_*$; $w_i' = w_i - w_*$. Let o be the set of fixed points through the deformation of point sets. Then, we have:

$$G(u) = \sum_j \Psi_j w_j' + w \quad (20)$$

Let $\Psi_j = (u - o_*) (\sum_i o_i'^\Phi \omega_i o_i')^{-1} \omega_j o_j'^\Phi$ be the scalar of image deformation speed determined by u and o ; u_{f1} and u_{f2} two angles on the horizontal direction in the region of pixel u ,

respectively; u_{f1}^* and u_{f2}^* be the two angles after the transformation, respectively; u_{f1} and u_{f2} be two angles on the vertical direction in the region of pixel u , respectively; u_{f1}^* and u_{f2}^* be the two angles after the transformation, respectively; ω_f and ω_u be the constrain weights in the horizontal direction and the vertical direction, respectively. To enhance the local structural constraint of the image series, the following constraint can be added:

$$\begin{aligned} & \underset{G}{\operatorname{argmin}} \sum_i u_i |G_u(o_i) - w_i|^2 \\ & + \omega_f |G_u(u) - \nu u_{f1}^* - (1 - \nu) u_{f2}^*|^2 \\ & + \omega_u |G_u(u) - \xi u_{u1}^* - (1 - \xi) u_{f2}^*|^2 \end{aligned} \quad (21)$$

where, ν satisfies the equation $u = \nu u_{f1} + (1 - \nu) u_{f2}$; ξ satisfies the equation $u = \xi u_{u1} + (1 - \xi) u_{u2}$. The desired transformation model can be obtained by solving the optimal result of formula (21) with moving least squares.

4. EXPERIMENTS AND RESULTS ANALYSIS

Based on the extracted image series features, the registration firstly relies on Harris corner detector to identify the corners, and then extracts the feature maps, according to the image series feature points outputted by the improved VGG16. Further, the feature maps are imported to the spatial transformation network for pre-registration of image series. Finally, the pre-registration results are deformed, restored, and resampled by region constraint-based moving least squares.

This paper compares image series pre-registration, the fully-connected properties of CNNs with different layers, and composite convolution features of four different kernel sizes. Tables 1 and 2 present the features on different layers of the image series extracted by original and improved VGG16, respectively. It can be observed that the image series registration features extracted by the second fully-connected layer were better than those of any other layer, as suggested by the relatively small RMSE and number of correctly matched points. In terms of the composite convolution features, CONV1-3 clearly outperformed CONV4, indicating that the kernel size directly affects the registration effect of composite convolution features.

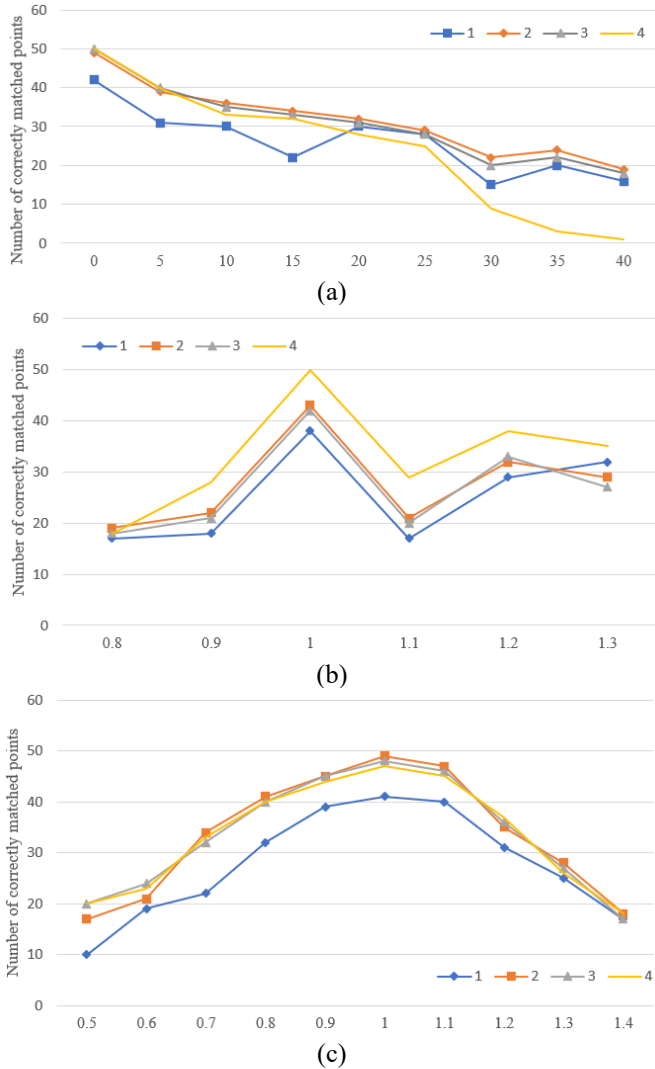
Table 1. Improved image registration precision

Features		FC1	FC2	CONV1	CONV2	CONV3	CONV4
Series 1	RMSE	0.0417	0.0412	0.0423	0.0426	0.0411	0.0462
	Jackknife						
	RMSE	0.0465	0.0454	0.0445	0.0418	0.0434	0.0459
	Number of correctly matched points	71	73	75	74	71	68
Series 2	RMSE	0.0814	0.0794	0.0815	0.0809	0.0827	0.0859
	Jackknife						
	RMSE	0.0919	0.0905	0.0911	0.0914	0.0906	0.0915
	Number of correctly matched points	33	35	29	24	28	26
Series 3	RMSE	0.0284	0.0275	0.0275	0.0299	0.0268	0.0319
	Jackknife						
	RMSE	0.0371	0.0329	0.0383	0.0348	0.0362	0.0418
	Number of correctly matched points	47	52	51	54	48	45

Note: RMSE is short for root-mean-square-error. The same below

Table 2. Image registration precision before the improvement

Features		FC1	FC2	CONV1	CONV2	CONV3	CONV4
Series 1	RMSE	0.0436	0.0428	0.0449	0.0428	0.0449	0.0529
	Jackknife	0.0482	0.0468	0.0493	0.0427	0.0468	0.0472
	RMSE						
	Number of correctly matched points	67	72	74	78	74	21
Series 2	RMSE	0.0828	0.0836	0.0861	0.0827	0.0818	0.0952
	Jackknife	0.0927	0.0948	0.0937	0.0917	0.0928	0.0962
	RMSE						
	Number of correctly matched points	25	22	26	25	28	21
Series 3	RMSE	0.0372	0.0362	0.0372	0.0362	0.0372	0.0419
	Jackknife	0.0327	0.0384	0.0327	0.0372	0.0358	0.0443
	RMSE						
	Number of correctly matched points	45	48	53	51	54	48

**Figure 7.** Image series matching quantities under different transformations

Comparative experiments were designed based on different transformations (e.g., rotation, scaling, and brightness change) to further evaluate the robustness of Harris corner detector in detecting angles, and the proposed region-constrained moving least squares facing different various transforms on the image series. The rotation angle was changed by an interval of 5° , and the scaling and brightness change was designed with an interval of 0.1 (unit: pixel and lumen). The high-performance layers of CONV1-3 and FC2 were tested. The number of

correctly matched points were counted after different transformations. The results (Figure 7) show that the orange curve, which represents the features outputted by FC2, outshined all the other curves under different transformations, while CONV3 achieved the best effect in scaling.

5. CONCLUSIONS

This paper mainly investigates the AI registration of image series with various features. Specifically, the Harris corner detector was called to extract the corners of multi-source image series, and the original and improved algorithms were described in details. After that, VGG16 was improved to extract the features of multi-source image series. Then, the spatial transformation network was employed for pre-registration of the image series, and the region-constrained moving least squares was used to deform and restore the image series. Through experiments, the image registration precisions of the original and improved VGG16 were compared, the high-performance layers of CONV1-3 and FC2 were tested, and the number of correctly matched points were counted after different transformations. The results show that the F2 features of our model work excellently under different transformations, and our model achieves a high effectiveness.

ACKNOWLEDGEMENTS

Project for Science and Technology Research of Jilin Province Education Department of China (Grant No.: JJKH20210696KJ).

REFERENCES

- [1] Degen, J., Modersitzki, J., Heinrich, M.P. (2015). Dimensionality reduction of medical image descriptors for multimodal image registration. *Current Directions in Biomedical Engineering*, 1(1): 201-205. <https://doi.org/10.1515/cdbme-2015-0051>
- [2] Palanivel, D.A., Natarajan, S., Gopalakrishnan, S. (2019). Mutifractals based multimodal 3D image registration. *Biomedical Signal Processing and Control*, 47: 126-136. <https://doi.org/10.1016/j.bspc.2018.08.015>
- [3] Ahmad, S., Khan, M.F. (2015). Deformable image registration based on elastodynamics. *Machine Vision*

- and Applications, 26(5): 689-710. <https://doi.org/10.1007/s00138-015-0690-1>
- [4] Lan, S., Guo, Z., You, J. (2019). Non-rigid medical image registration using image field in Demons algorithm. *Pattern Recognition Letters*, 125: 98-104. <https://doi.org/10.1016/j.patrec.2019.04.006>
- [5] Yigitsoy, M., Navab, N. (2013). Structure propagation for image registration. *IEEE Transactions on Medical Imaging*, 32(9): 1657-1670. <https://doi.org/10.1109/TMI.2013.2263151>
- [6] Wang, C., Goatman, K.A., Boardman, J.P., Beveridge, E.L., Newby, D.E., Semple, S.I. (2019). Distance oriented particle swarm optimizer for brain image registration. *IEEE Access*, 7: 56016-56027. <https://doi.org/10.1109/ACCESS.2019.2907769>
- [7] Costa, P.B., Marques, A., Baldner, F.D.O., Leta, F.R. (2013). Line scale measurement using image registration. *International Journal of Metrology and Quality Engineering*, 4(2): 121-125.
- [8] Guyader, J.M., Huizinga, W., Fortunati, V., Poot, D.H., Veenland, J.F., Paulides, M.M., Klein, S. (2018). Groupwise multichannel image registration. *IEEE Journal of Biomedical and Health Informatics*, 23(3): 1171-1180. <https://doi.org/10.1109/JBHI.2018.2844361>
- [9] Tang, S., Fang, F., Zhang, G. (2015). Variational approach for multi-source image fusion. *IET Image Processing*, 9(2): 134-141.
- [10] Wu, H., An, D., Zhu, X., Zhang, Z., Fan, G., Hua, Z. (2021). Multi-source material image optimized selection based multi-option composition. *Image and Vision Computing*, 107: 104123. <https://doi.org/10.1016/j.imavis.2021.104123>
- [11] Yang, Z., Chen, Y., Le, Z., Fan, F., Pan, E. (2019). Multi-source medical image fusion based on Wasserstein generative adversarial networks. *IEEE Access*, 7: 175947-175958. <https://doi.org/10.1109/ACCESS.2019.2955382>
- [12] Solav, D., Moerman, K.M., Jaeger, A.M., Genovese, K., Herr, H.M. (2018). MultiDIC: An open-source toolbox for multi-view 3D digital image correlation. *IEEE Access*, 6: 30520-30535. <https://doi.org/10.1109/ACCESS.2018.2843725>
- [13] Zhang, H., Ren, P. (2017). Game theoretic hypergraph matching for multi-source image correspondences. *Pattern Recognition Letters*, 87: 87-95. <https://doi.org/10.1016/j.patrec.2016.07.011>
- [14] Ma, G., Piao, Y., Li, B. (2017). Research on multi-source image fusion technology in haze environment. In *LIDAR Imaging Detection and Target Recognition 2017*, 10605: 1060531. <https://doi.org/10.1117/12.2295089>
- [15] Li, J., Peng, Z. (2015). Multi-source image fusion algorithm based on cellular neural networks with genetic algorithm. *Optik*, 126(24): 5230-5236. <https://doi.org/10.1016/j.ijleo.2015.09.187>
- [16] Zhang, Y., Prasad, S. (2014). Locality preserving composite kernel feature extraction for multi-source geospatial image analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(3): 1385-1392. <https://doi.org/10.1109/JSTARS.2014.2348537>
- [17] Guo, Y.L., Wang, Y.B., Li, Y.M., Wang, Q., Zhu, L., Lü, H. (2015). Research of multi-source water remote image fusion algorithm based on bio-optical model. *Guangxue Xuebao/Acta Optica Sinica*, April 10, 2015, 35(4).
- [18] Xu, G., Wu, Q., Cheng, Y., Yan, F., Li, Z., Yu, Q. (2021). A robust deformed image matching method for multi-source image matching. *Infrared Physics & Technology*, 115: 103691. <https://doi.org/10.1016/j.infrared.2021.103691>
- [19] Yan, Z., Wicaksana, J., Wang, Z., Yang, X., Cheng, K.T. (2020). Variation-Aware federated learning with multi-source decentralized medical image data. *IEEE Journal of Biomedical and Health Informatics*, 25(7): 2615-2628. <https://doi.org/10.1109/JBHI.2020.3040015>
- [20] Wang, Y., Lu, T., Zhang, Y., Wu, Y. (2021). Multi-scale Self-calibrated Network for Image Light Source Transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 252-259.
- [21] Dolly, D.R.J., Peter, J.D., Bala, G.J., Jagannath, D.J. (2020). Image fusion for stabilized medical video sequence using multimodal parametric registration. *Pattern Recognition Letters*, 135: 390-401. <https://doi.org/10.1016/j.patrec.2020.04.001>
- [22] Mingjun, W., Zhen, Z., Liangliang, Z., Guodong, S. (2012). A fast and memory-saving algorithm for sequence-based image registration. *High Power Laser and Particle Beams*, 24(5): 1038-1042. <https://doi.org/10.3788/HPLPB20122405.1038>
- [23] Chen, W.L., He, X.H., Song, H.Y., Wu, W. (2012). An improved sequence-to-sequence alignment method combined with feature-based image registration algorithm. *Journal of Information Science and Engineering*, 28(3): 617-630.
- [24] Kirby, R., Whitaker, R. (2016). A novel automated method for doing registration and 3D reconstruction from multi-modal RGB/IR image sequences. In *Infrared Sensors, Devices, and Applications VI*, 9974: 997400. <https://doi.org/10.1117/12.2237165>
- [25] Fan, B., Du, Y., Tang, Y. (2011). Efficient registration algorithm for UAV image sequence. In *2011 IEEE International Conference on Information and Automation*, 111-116. <https://doi.org/10.1109/ICINFA.2011.5948972>
- [26] Li, Y., Shi, D.L., Bu, F.J. (2019). Automatic recognition of rock images based on convolutional neural network and discrete cosine transform. *Traitement du Signal*, 36(5): 463-469. <https://doi.org/10.18280/ts.360512>
- [27] Zhang, W.L., Li, X.W., Song, Q.X., Lu, W. (2020). A face detection method based on image processing and improved adaptive boosting algorithm. *Traitement du Signal*, 37(3): 395-403. <https://doi.org/10.18280/ts.370306>