Target Registration Based on Fusing Features of Visible and Two Wave Bands Infrared Images

Junhua Yan[1,2(✉)], Kai Su[2], Xuyang Cai[2], Tianxia Xie[2], Yin Zhang[1,2], and Kun Zhang[2]

[1] Key Laboratory of Space Photoelectric Detection and Perception, Nanjing University of Aeronautics and Astronautics, Ministry of Industry and Information Technology, No. 29 Jiangjun Street, Nanjing 211106, China
yjh9758@126.com

[2] College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, Jiangsu, China

Abstract. In order to register the same target in images from different sources to improve the accuracy of target recognition of multi-source images, based on the principle that the same target has the highest similarity among targets in these images, this paper proposes a new target registration algorithm by fusing features of Visible (VIS), Long Wave Infrared (LWIR) and Middle Wave Infrared (MWIR) images, which registers the same target in these images by calculating the targets similarity in different source images. Firstly, the similarity between targets in LWIR and MWIR images is calculated by using the improved structural similarity. Then, the similarity between targets in VIS and LWIR images is calculated by using Hu invariant moment feature and cosine similarity. Finally, the similarity among targets in VIS, MWIR and LWIR images is calculated by fusing these two kinds of target similarity, so that target registration of these three-source images is realized. Experimental results show that the proposed algorithm has high correct rate and accuracy of target registration. Specifically, the correct rate of target registration is 83.87% and the accuracy of target registration is higher than 0.95.

Keywords: Target registration · Fusion · VIS image · LWIR image · MWIR image · Improved structural similarity

1 Introduction

The expansion of human activities in time and space demands recognition of interested targets from complex background under all weather conditions, hence multi-source images are needed to realize this demand. In order to improve the accuracy of target recognition of multi-source images, the same target in these images must be registered firstly. Because of different imaging principles, images from different sources have different image features and the same target in these images has different target features. VIS images have rich color information and texture information to capture the details of the target. However, adverse weather conditions render VIS images

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2021
Published by Springer Nature Switzerland AG 2021. All Rights Reserved
S. Shi et al. (Eds.): AICON 2020, LNICST 356, pp. 465–483, 2021.
https://doi.org/10.1007/978-3-030-69066-3_41

ineffective. Infrared images, by contrast, are not affected by the weather. Not many details are captured in infrared images, yet the edge information of the target, which provides the complete structural information about the target, is well preserved. Specifically, MWIR images are richer in texture information than LWIR images, especially in high thermal radiation regions. LWIR images, however, are brighter than MWIR images. Although the same target has different target features in images from different sources, it has the highest similarity among targets in these images. In order to recognize targets effectively under all weather conditions and at all times, the target features of the visible and two wave bands infrared images are fused in this paper based on the targets similarity in different source images, realizing target registration of multi-source images.

J. Ma et al. [1] realized the registration of corresponding regions in the VIS and LWIR images. This method combines edge enhancement and normalized correlation coefficient, and the registration rate reaches 83.33%. This algorithm highlights the complete edge information of LWIR images through edge enhancement. However, the normalized correlation coefficient is only related to the gray information of the image. Moreover, the different imaging principles of VIS and LWIR images cause significant differences in the gray of these images, which renders the algorithm ineffective when the image is complex. J. Jiang et al. [2] registered the VIS and LWIR images based on wavelet transform and mutual information, with a registration rate of 80%. This method extracts the complete edge information of VIS and LWIR images with wavelet transform. However, the computation of mutual information is entirely based on the statistical information of the image gray, leaving out the spatial information of the image gray. Thus, similarly, this algorithm is ineffective when the image is complex. L. L. Mao et al. [3] combined the mutual information and the gradient information to realize the registration of VIS and LWIR images, and the registration rate is higher than those of registration algorithms which are only based on mutual information. This algorithm uses the gradient information to represent the complete edge information of VIS and LWIR images. However, this algorithm do not have high robustness, because the mutual information is only related to the statistical information of the image gray, and the spatial distribution of the image gray is ignored. Z. Li [4] put forward an algorithm to register VIS and LWIR images using the normalized covariance correlation coefficient. Although the normalized covariance correlation coefficient is strongly adaptable to gray transformations between images, it belongs to gray correlation. Therefore, this algorithm is not suitable for a variety of images from different sources. G.A. Bilodeau et al. [5] used the histogram of oriented gradient to register VIS and LWIR images. The registration rate is around 50% when the registration window is small and can reach more than 90% when the registration window becomes larger. Although the histogram of oriented gradient can represent the common edge information of VIS and LWIR images, this algorithm is not robust to the sizes of registration windows. M. An et al. [6] presented a robust image registration and localization algorithm based on SURF features to register remote sensing video images. This algorithm can compute camera poses under different conditions of scale, viewpoint and rotation so as to precisely localize object's position. However, the algorithm almost based on the land-marks, it only works when landmarks are available in the scene. J.Y. Ma [7] proposed a regularized Gaussian fields criterion for non-rigid registration of visible and infrared face images, which represents an image by its edge map and align the edge maps by a robust criterion with a non-rigid model. This algorithm is suitable for non-rigid bodies such as face images, but it is not robust to images lacking of rich edge information. Y. Zhuang [8] registered visible and infrared face images based on mutual information, which combines the PSO algorithm and Powell search method to find the most appropriate registration parameters. However, mutual information is entirely based on the statistical information of the image gray so this method is not suitable for complex images.

The above mentioned registrations algorithms of multi-source images make good use of the complete edge information of VIS and LWIR images. However, they only use the statistical information of the gray or the edge orientation to characterize the edge information of these images, and do not utilize the complete shape and structural information. Hence they are not suitable for the registration of a variety of complex images from different sources. The Hu invariant moment feature is a typical shape feature that can represent the shape information of the target in the image. J.F. Dou et al. [9] proposed a registration algorithm of VIS and LWIR images, in which the Hu invariant moment feature of the neighboring area around every feature point is used to match feature points. This algorithm makes stable use of the shape feature of the neighboring area around every feature point, and has a good registration performance in the presence of changes in rotation, scale, translation, field, etc. X. Wen et al. [10] put forward an improved structural similarity algorithm. For the overall structural feature of the target, it uses the images and the corresponding edge images of different scales in Gaussian scale space to calculate the multi-scale edge structural similarity, realizing the target registration of two wave bands infrared images. The correct rate and accuracy of the target registration of this algorithm is high. F. Wu [11] presented an image registration method improving the precision of visible and infrared (VIS/IR) image registration based on visual salient (VS) feature detector. This detector can detect the VS correspondences between images with higher repeatability score. However, this method is not applicable to target with big scale change. St-Charles et al. [12] presented an online multimodal video registration method that relies on the matching of shape contours to estimate the parameters of a planar transformation model. They used foreground-background segmentation on each video frame to obtain shape contours from targets present in the scene, then described and matched contour points using the iterative shape context approach. However, the algorithm is not suitable for targets with approximate shape. Y.J. Chen [13] presented an image registration method for visible and infrared images based on stable region features and edginess, which uses Zernike moments to describe salient region features for a coarse registration, and uses an entropy optimal process based on edginess to refine the registration to achieve a more accurate result. However, Zernike invariant moments are effective for images in which the target shape is dominant, and are less effective for texture-rich images.

Hence, the edge information, shape information and structure information in VIS, LWIR and MWIR images are fully utilized in this paper. Based on the principle that the same target has the highest similarity among targets in images from different sources, target features in VIS, LWIR and MWIR images are fused by calculating the targets similarity, which improves the registration accuracy of the same target in these images. Firstly, the similarity between targets in LWIR and MWIR images is calculated by using the improved structural similarity. Then, the similarity between targets in VIS and LWIR images is calculated by using Hu invariant moment feature and cosine similarity, which utilizes the complete shape information of VIS and LWIR images. Finally, the similarity among targets in VIS, MWIR and LWIR images is calculated by fusing these two kinds of target similarity, so that target registration based on fusing features of VIS and two wave bands infrared images is realized. In this paper, the proposed algorithm is tested on multiple sets of VIS, MWIR and LWIR images, and the superiority of the proposed algorithm is verified.

2 The Target Registration Algorithm Based on Fusing Features of VIS and Two Wave Bands Infrared Images

The framework of the target registration algorithm is shown in Fig. 1.

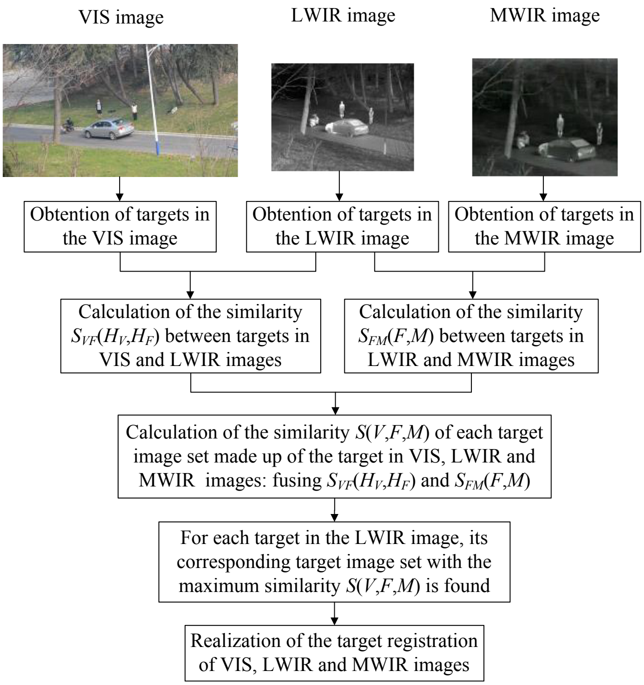


Fig. 1. The framework of the target registration algorithm based on fusing features of VIS and two wave bands infrared images

Based on the principle that the same target has the highest similarity among targets in images from different sources, target features in VIS, LWIR and MWIR images are fused by calculating the targets similarity, which realizes the registration of the same target in these images. Firstly, since LWIR and MWIR images contain rich edge information and complete structure information, the similarity $S_{FM}(F,M)$ between targets in LWIR and MWIR images is calculated by using the improved structural similarity; since VIS and LWIR images contain rich edge information and complete geometric shapes information, the similarity $S_{VF}(H_V,H_F)$ between targets in VIS and LWIR images is calculated by using the Hu invariant moment feature and cosine similarity. Then, the similarity $S(V,F,M)$ among targets in VIS, LWIR and MWIR images is calculated by fusing $S_{FM}(F,M)$ and $S_{VF}(H_V,H_F)$ in the same proportion. Finally, for each target in LWIR image, the corresponding target image set with the maximum value of similarity $S(V,F,M)$ is identified by searching, which realizes the target registration of VIS, LWIR and MWIR images.

3 Calculation of the Similarity Among Targets in VIS and Two Wave Bands Infrared Images

3.1 Calculation of the Similarity Between Targets in LWIR and MWIR Images

LWIR and MWIR images contain rich edge information and complete structure information [14]. The improved structural similarity uses these images and their edge images of different scales in Gaussian scale space to calculate the multi-scale structural similarity. It combines the complete edge information with the structural information of the target in LWIR and MWIR images [10]. Therefore, the similarity between targets in LWIR and MWIR images is calculated by using the improved structural similarity. The specific computational formula is as follows:

$$S_{FM}(F,M) = [l_{MSG}(F,M)]^{\alpha}\cdot[c_{MSG}(F,M)]^{\beta}\cdot[s_{MSeG}(F,M)]^{\gamma},\ 0<\alpha,\beta,\gamma\leq 1 \quad (1)$$

where $l_{MSG}(F,M) = [l(F_N,M_N)]^{w_N}$, representing the value of multi-scale luminance comparison.

$c_{MSG}(F,M) = \prod_{i=1}^{N}[c(F_i,M_i)]^{w_i}$, representing the value of multi-scale contrast comparison.

$s_{MSeG}(F,M) = \prod_{i=1}^{N}[s_e(F_i,M_i)]^{w_i}$, representing the value of multi-scale structure comparison.

Where $F_i = F*G(x,y,t_i)$, $M_i = M*G(x,y,t_i)$, $i = 1,2,\ldots,N$, N is the number of scales, F_i and M_i represent the Gaussian scale space of LWIR image F and MWIR image M respectively. Here, $N = 5$, and the values of w_i ($i = 1,2,\ldots,5$) are 0.0448, 0.2856, 0.3001, 0.2326, 0.1333 respectively.

3.2 Calculation of the Similarity Between Targets in VIS and LWIR Images

VIS and LWIR images contain rich edge information and complete geometric shapes [15]. However, camera motion causes changes of scale, translation and rotation between images. Hu invariant moment feature can describe the shape information of the image, and it is invariant to scale, translation and rotation [6]. Hence, it can serve as the effective feature of the target in VIS and LWIR images to realize target registration. Therefore, Hu invariant moment feature vector of 7 dimensions is extracted in this paper. The computational formula is as follows:

$$Hu_0 = \eta_{20} + \eta_{02}; \tag{2}$$

$$Hu_1 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2; \tag{3}$$

$$Hu_2 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2; \tag{4}$$

$$Hu_3 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2; \tag{5}$$

$$\begin{aligned} Hu_4 = {} & (\eta_{30} - 3\eta_{21})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \cdot [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]; \end{aligned} \tag{6}$$

$$Hu_5 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}); \tag{7}$$

$$\begin{aligned} Hu_6 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03}) \cdot [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]; \end{aligned} \tag{8}$$

Where η_{pq} (p, q = 0,1,2,3) is the normalized central moment of (p + q) order of the image.

After extracting the Hu invariant moment feature of the target in VIS and LWIR images respectively, cosine similarity [16] is used to measure the similarity between targets in VIS and LWIR images. The computational formula of cosine similarity is as follows:

$$S_{VF}(H_V, H_F) = \frac{\sum_i H_V(i) \cdot H_F(i)}{\sqrt{\sum_i (H_V(i))^2 \cdot \sum_i (H_F(i))^2}} \tag{9}$$

Where HV and H_F represent the Hu invariant moment feature vector of the target in VIS and LWIR images respectively; i represents the dimension of the Hu invariant moment feature vector, and i = 1,2,…,7.

3.3 Calculation of the Similarity Among Targets in VIS and Two Wave Bands Infrared Images

Based on the similarity among targets in images from different sources, target features in VIS, LWIR and MWIR images are fused to realize the registration of the same target in these images. Firstly, the similarity $S_{VF}(H_V, H_F)$ between targets in VIS and LWIR images is calculated by using the Hu invariant moment feature and cosine similarity; the similarity $S_{FM}(F, M)$ between targets in LWIR and MWIR images is calculated by using the improved structural similarity. Then, the similarity $S(V, F, M)$ among targets in VIS, LWIR and MWIR images is calculated by fusing $S_{FM}(F, M)$ and $S_{VF}(H_V, H_F)$ in the same proportion. The calculation formula is as follows:

$$S(V, F, M) = 0.5 \times S_{VF}(H_V, H_F) + 0.5 \times S_{FM}(F, M) \tag{10}$$

4 Realization of Target Registration of VIS and Two Wave Bands Infrared Images

The flow chart of target registration of VIS and two wave bands infrared images is shown in Fig. 2.

In Fig. 2, N_0 represents the number of targets in the VIS image; N_1 represents the number of targets in the LWIR image; N_2 represents the number of targets in the MWIR image; n_V represents the index of the target in the VIS image; n_M represents the index of the target in the MWIR image; n_F represents the index of the target in the LWIR image.

Realization steps:

(1) Target regions of VIS, LWIR and MWIR images are extracted.
(2) The similarity between targets in VIS and LWIR images is calculated by using Hu invariant moment feature and cosine similarity.
(3) The similarity between targets in LWIR and MWIR images is calculated by using the improved structural similarity.
(4) The similarity among targets in VIS, LWIR and MWIR images is calculated by fusing these two kinds of target similarity abovementioned.
(5) For the target n_F (the initial value of n_F is 1) in the LWIR image, the similarity between the target n_F and the target n_V (the initial value of n_V is 1) as well as the similarity between the target n_F and the target n_M (the initial value of n_M is 1) are calculated. Then, determine the similarity of this target image set is the maximum similarity of the target n_F or not. If yes, registration of the target n_F with targets in VIS and MWIR images is realized; if no, the value of n_M is increased by 1 and the process aforementioned is repeated. If the maximum similarity of the target n_F is not found after all targets in the MWIR image are tested, the value of n_V is increased by 1, the value of n_M is reset to 1 and the process aforementioned is repeated.
(6) Determine whether n_F is the terminate value or not. If yes, the program ends; if no, the value of n_F is increased by 1, the values of n_V and n_M is reset to 1, and the step (5) is repeated.

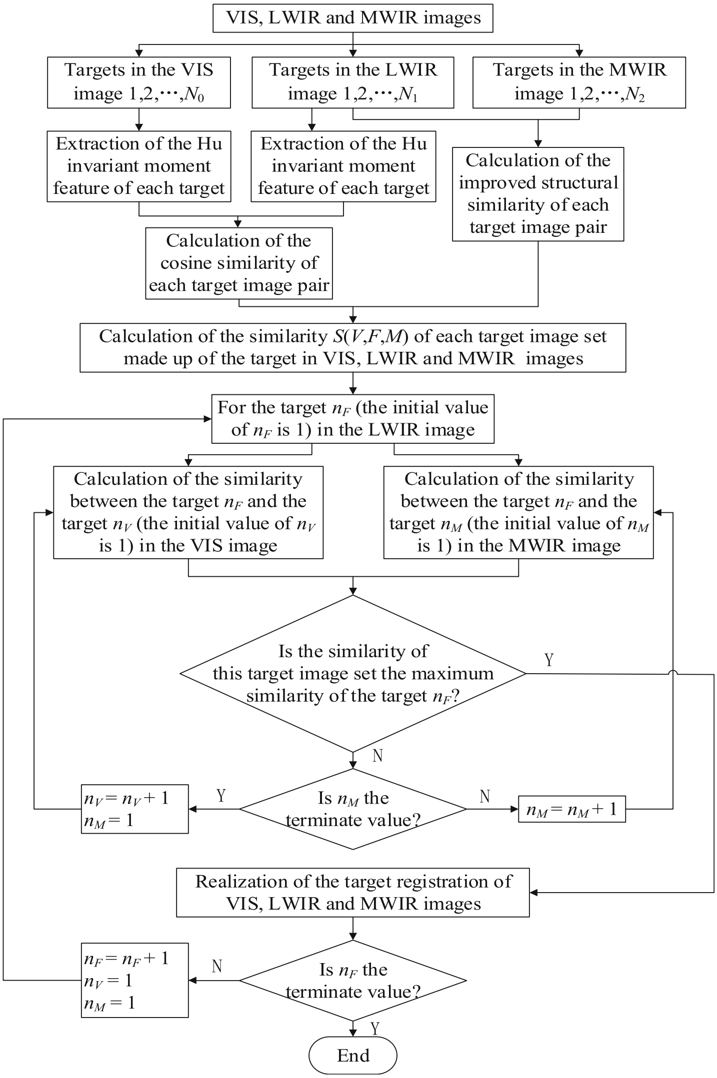


Fig. 2. The flow chart of target registration of VIS and two wave bands infrared images

5 Experimental Results and Discussion

The experimental images in this paper are VIS, LWIR and MWIR images of the same scene. With extensive research, no open data set was found to provide all these three types of source images for a same scene. Therefore, 9 sets of experimental images were taken in the field. VIS images were taken by a digital single lens reflex camera, and the resolution is 1920 × 1080 pixels; LWIR images were taken by a handheld infrared thermal imager with a single eyepiece (type specification: IR518, working band: 8–14 μm), and the resolution is 384 × 288 pixels; MWIR images were taken by a cooled thermal imager (type specification: IR300, working band: 3–5 μm), and the resolution is 320 × 256 pixels. Due to the long shooting distance, the targets occupy a small percentage of pixels in the entire images, which shows that the targets are small.

The proposed algorithm in this paper is named TRA_VFM algorithm. The proposed algorithm and the target registration algorithms of VIS and LWIR images mentioned in literatures [1–5] are applied to 9 sets of VIS, LWIR and MWIR images. The experimental purpose is to compare the accuracy of target registration of various registration algorithms, in order to verify the superiority of the proposed algorithm. Each image in experimental image sets 1–5 includes 3 targets while each image in experimental image sets 6–9 includes 4 targets. Images containing the same target with the same attitude in VIS, MWIR and LWIR images are considered as a set of target images, adding up to 31 sets of target images; images containing the same target with the same attitude in VIS and LWIR images are considered as a pair of target images, adding up to 31 pairs of target images.

All experiments were carried out on ASUS's G11 flight fortress desktop computer with the Intel Core i7–7700 processor and 8G memory. The operating system is Win7 and the experimental software isVS2010 + OpenCV2.4.6.

5.1 Target Registration Experiment of VIS and Two Wave Bands Infrared Images

The six registration algorithms are TRA_VFM algorithm, NorCroCor algorithm [1], NormMI algorithm [2], NormMI_Gradient algorithm [3], NorCovCor algorithm [4] and HOG algorithm [5] respectively. Experimental results are shown in Fig. 3, Fig. 4, and Fig. 5. Targets in boxes of the same color (red, green, yellow, or pink) represent the same target. Specifically, targets in red boxes, green boxes, yellow boxes and pink boxes are target 1, target 2, target 3 and target 4 respectively. For each image set processed by the TRA_VFM algorithm, the upper left image is the LWIR image, the upper right image is the MWIR image and the lower image is the VIS image. For each image pair processed by the other 5 algorithms, the upper image is the LWIR image and the lower image is the VIS image.

There are changes in image background, target attitude, shooting angle, etc. among the 9 image sets. Experimental results shown in Fig. 3, Fig. 4 and Fig. 5 indicate that the correct rate of target registration of the TRA_VFM algorithm is higher than that of the other 5 target registration algorithms. Moreover, the TRA_VFM algorithm realizes the registration of the target from three different image sources, which are VIS, LWIR and MWIR images. However, the other 5 algorithms can only realize target registration in VIS and LWIR images. For the TRA_VFM algorithm, target 1 and target 2 are registered erroneously in image set 2, target 1 is registered erroneously in image sets 3 and 5, and target 2 is registered erroneously in image set 6. For the NorCroCor algorithm, the NormMI algorithm and the HOG algorithm, some targets are registered erroneously in image sets 1 to 9. For the NormMI_Gradient algorithm, all targets are registered correctly in image sets 1, 3, and 9, but some targets are registered erroneously in the rest of the image sets. For the NorCovCor algorithm, all targets are

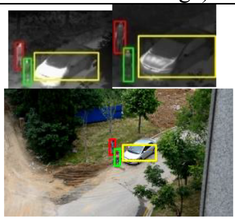
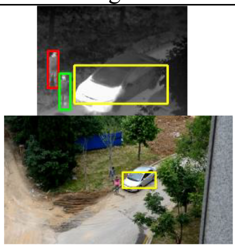
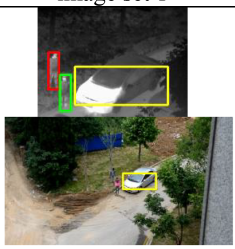
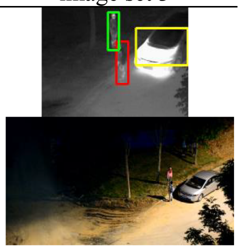
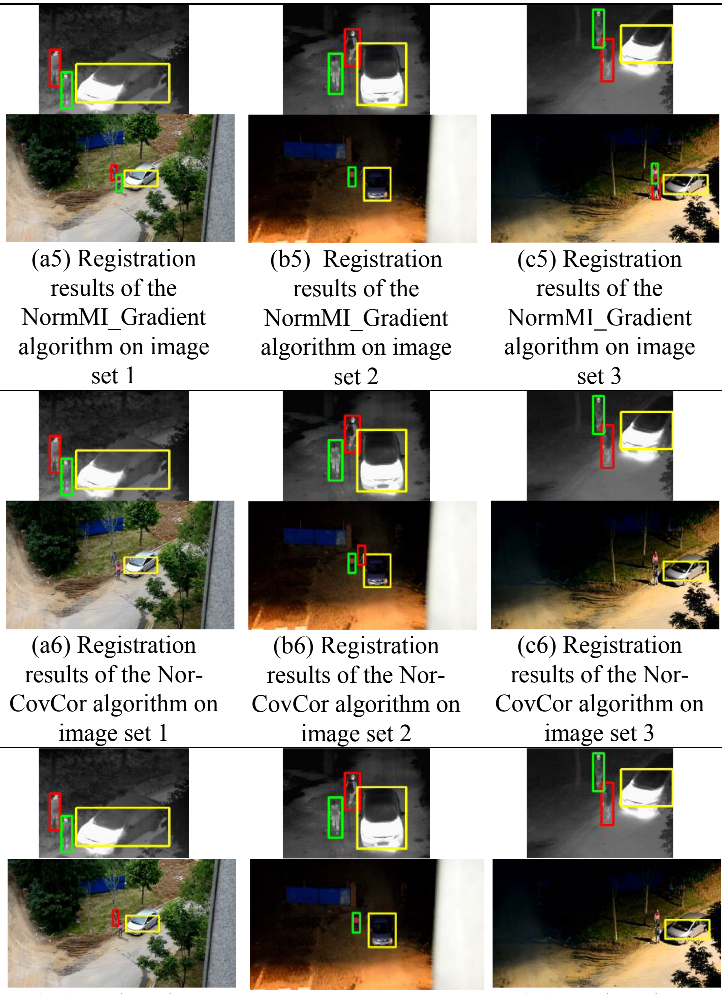
	Image set 1	Image set 2	Image set 3
Input image	 (a1) Original image set 1 (upper-left: LWIR image, upper-right: MWIR image, bottom: VIS image)	 (b1) Original image set 2(upper-left: LWIR image, upper-right: MWIR image, bottom: VIS image)	 (c1) Original image set 3 (upper-left: LWIR image, upper-right: MWIR image, bottom: VIS image)
TRA_VFM	 (a2) Registration results of the TRA_VFM algorithm on image set 1	 (b2) Registration results of the TRA_VFM algorithm on image set 2	 (c2) Registration results of the TRA_VFM algorithm on image set 3
Nor-CroCor	 (a3) Registration results of the Nor-CroCor algorithm on image set 1	 (b3) Registration results of the Nor-CroCor algorithm on image set 2	 (c3) Registration results of the Nor-CroCor algorithm on image set 3
NormMI	 (a4) Registration results of the NormMI algorithm on image set 1	 (b4) Registration results of the NormMI algorithm on image set 2	 (c4) Registration results of the NormMI algorithm on image set 3

Fig. 3. Target registration results of six algorithms on image sets 1 to 3



NormMI_ Gradient

(a5) Registration results of the NormMI_Gradient algorithm on image set 1

(b5) Registration results of the NormMI_Gradient algorithm on image set 2

(c5) Registration results of the NormMI_Gradient algorithm on image set 3

Nor-CovCor

(a6) Registration results of the Nor-CovCor algorithm on image set 1

(b6) Registration results of the Nor-CovCor algorithm on image set 2

(c6) Registration results of the Nor-CovCor algorithm on image set 3

HOG

(a7) Registration results of the HOG algorithm on image set 1

(b7) Registration results of the HOG algorithm on image set 2

(c7) Registration results of the HOG algorithm on image set 3

Fig. 3. (*continued*)

registered correctly in image set 2, but some targets are registered erroneously in the rest of the image sets.

In order to compare the correct rates of target registration of the 6 tested target registration algorithm quantitatively, an evaluation index is put forward in this paper: the correct rate of target registration. For the TRA_VFM algorithm, the correct rate of target registration is defined as the ratio of the number of the target image sets registered correctly to the total number of target image sets; for the other 5 target registration algorithms, the correct rate of target registration is defined as the ratio of the number of the target image pairs registered correctly to the total number of target image pairs. The correct rate of target registration of these 6 target registration algorithms is shown in Table 1.

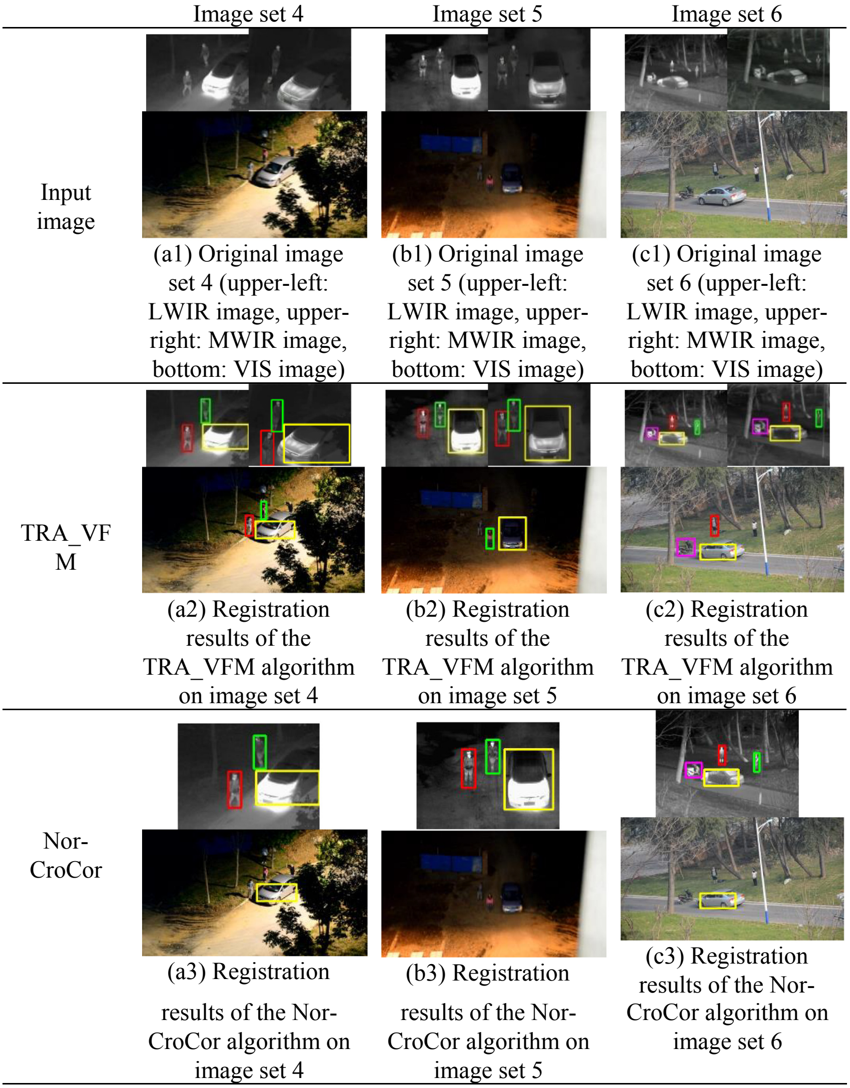


Fig. 4. Target registration results of six algorithms on image sets 4 to 6

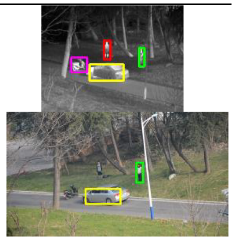
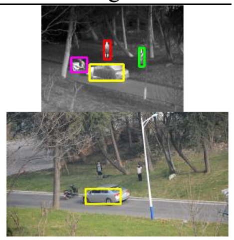
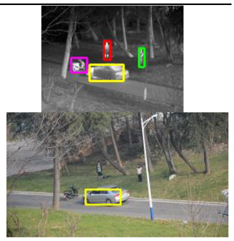
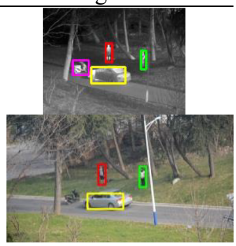
NormMI	 (a4) Registration results of the NormMI algorithm on image set 4	 (b4) Registration results of the NormMI algorithm on image set 5	 (c4) Registration results of the NormMI algorithm on image set 6
NormMI_ Gradient	 (a5) Registration results of the NormMI_Gradient algorithm on image set 4	 (b5) Registration results of the NormMI_Gradient algorithm on image set 5	 (c5) Registration results of the NormMI_Gradient algorithm on image set 6
Nor-CovCor	 (a6) Registration results of the Nor-CovCor algorithm on image set 4	 (b6) Registration results of the Nor-CovCor algorithm on image set 5	 (c6) Registration results of the Nor-CovCor algorithm on image set 6
HOG	 (a7) Registration results of the HOG algorithm on image set 4	 (b7) Registration results of the HOG algorithm on image set 5	 (c7) Registration results of the HOG algorithm on image set 6

Fig. 4. (*continued*)

Table 1 shows that the correct rate of target registration of the TRA_VFM algorithm is 83.87% and those of the other 5 algorithms are less than 62%. The results above prove that the TRA_VFM algorithm has a high correct rate of target registration that is much higher than the correct rates of the NorCroCor algorithm, the NormMI algorithm, the NormMI_Gradient algorithm, the NorCovCor algorithm and the HOG algorithm.

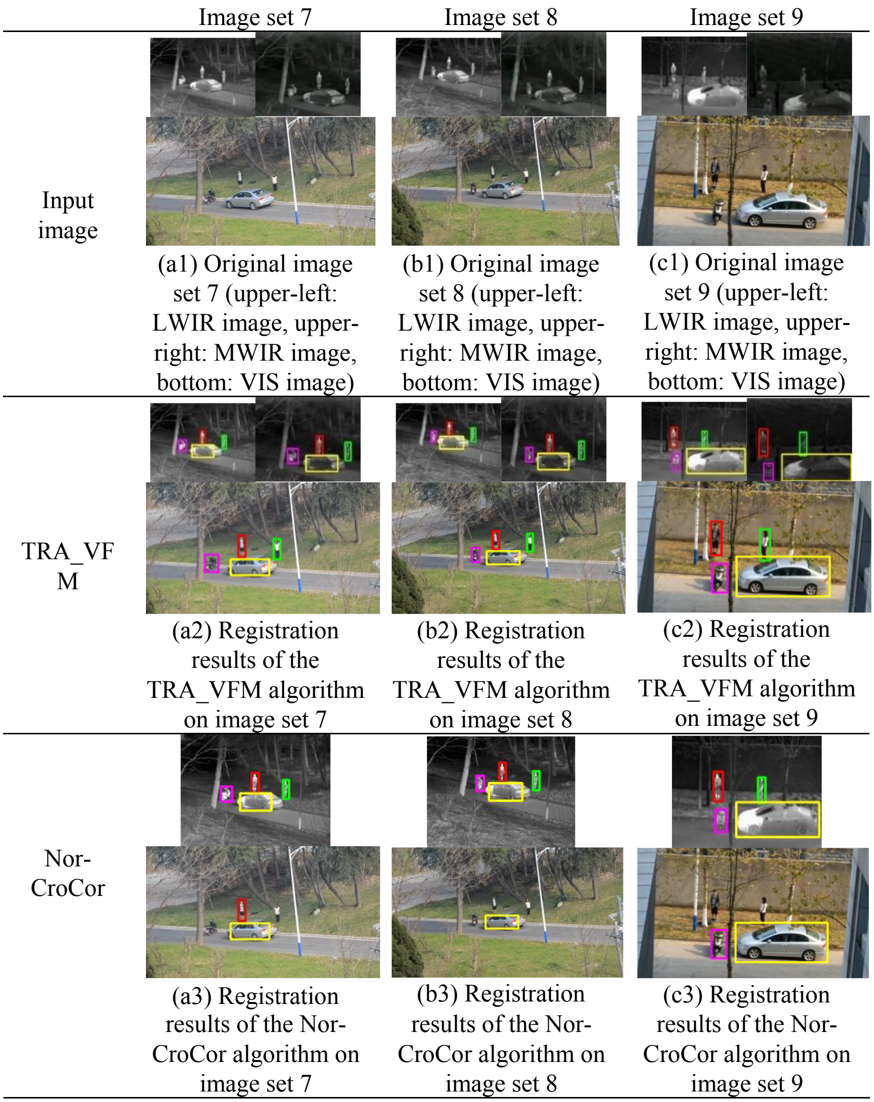


Fig. 5. Target registration results of six algorithms on image sets 7 to 9

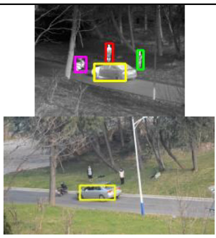
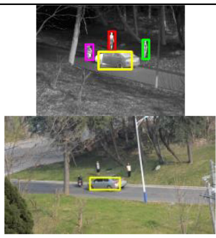
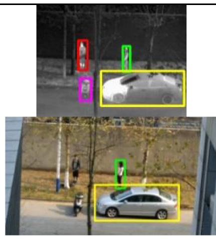
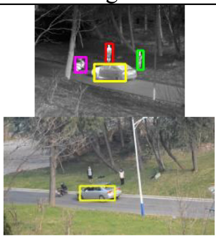
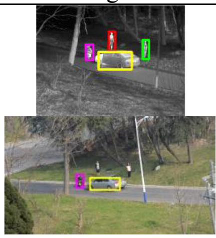
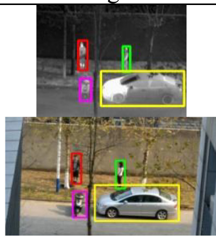
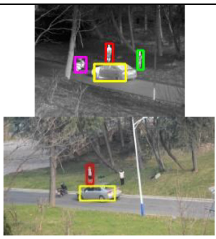
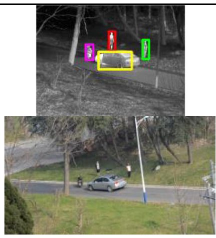
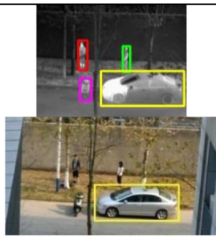
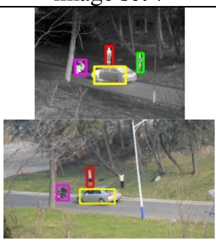
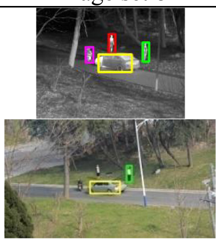
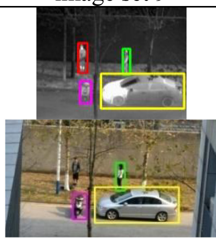
NormMI	 (a4) Registration results of the NormMI algorithm on image set 7	 (b4) Registration results of the NormMI algorithm on image set 8	 (c4) Registration results of the NormMI algorithm on image set 9
NormMI_ Gradient	 (a5) Registration results of the NormMI_Gradient algorithm on image set 7	 (b5) Registration results of the NormMI_Gradient algorithm on image set 8	 (c5) Registration results of the NormMI_Gradient algorithm on image set 9
Nor-CovCor	 (a6) Registration results of the Nor-CovCor algorithm on image set 7	 (b6) Registration results of the Nor-CovCor algorithm on image set 8	 (c6) Registration results of the Nor-CovCor algorithm on image set 9
HOG	 (a7) Registration results of the HOG algorithm on image set 7	 (b7) Registration results of the HOG algorithm on image set 8	 (c7) Registration results of the HOG algorithm on image set 9

Fig. 5. (*continued*)

Table 1. The correct rate of target registration of the 6 tested algorithms (%)

	TRA_VFM	NorCroCor	NormMI	NormMI_Gradient	NorCovCor	HOG
The correct rate of target registration	**83.87**	45.16	38.71	61.29	41.94	61.29

5.2 Accuracy of the Target Registration of VIS and Two Wave Bands Infrared Images

The TRA_VFM algorithm is compared with 3 algorithms which use similarity as the index of target registration on the accuracy of target registration: the NorCroCor algorithm, the NorCovCor algorithm and the HOG algorithm. The NormMI algorithm and the NormMI_Gradient algorithm use mutual information as the index of target registration. Hence it is not suitable to compare them with the TRA_VFM algorithm on the accuracy of target registration. The evaluation index put forward in this paper is the accuracy of target registration. For the TRA_VFM algorithm, the accuracy of target registration is defined as the similarity among targets registered correctly in VIS, LWIR and MWIR images; for the other 3 target registration algorithms, the accuracy of target registration is defined as the similarity between targets registered correctly in VIS and LWIR images.

The values of the similarity among or between targets registered correctly in 9 image sets are shown in Table 2. Among the similarity values registered by these 4 algorithms in every target image set or target image pair, every maximum value is marked in bold, and "###" indicates that the target image set or target image pair is registered incorrectly. In image sets 1–5, targets in each LWIR image are FO_1, FO_2 and FO_3, targets in each MWIR image are MO_1, MO_2 and MO_3, and targets in each VIS image areVO_1, VO_2 and VO_3. In image sets 6–9, targets in each LWIR image are FO_1, FO_2, FO_3 and FO_4, targets in each MWIR image are MO_1, MO_2, MO_3 and MO_4, and targets in each VIS image are VO_1, VO_2, VO_3 and VO_4. Besides, (FO_1, MO_1, VO_1), (FO_2, MO_2, VO_2), (FO_3, MO_3, VO_3), (FO_4, MO_4, VO_4) are the same target respectively.

Table 2. Similarity values for correctly registered targets in 9 image sets processed by 6 registration algorithms

Image set	Target set	TRA_VFM	Target pair	NorCroCor	NorCovCor	HOG
Image set 1	FO_1-MO_1-VO_1	**0.9995**	FO_1-MO_1	###	###	0.9882
	FO_2-MO_2-VO_2	**0.9997**	FO_2-MO_2	###	###	###
	FO_3-MO_3-VO_3	**0.9950**	FO_3-MO_3	0.8163	0.3213	0.9832
Image set 2	FO_1-MO_1-VO_1	###	FO_1-MO_1	**0.8542**	0.2552	###
	FO_2-MO_2-VO_2	###	FO_2-MO_2	0.8685	0.2464	**0.9939**
	FO_3-MO_3-VO_3	**0.9660**	FO_3-MO_3	###	0.5129	0.9163
Image set 3	FO_1-MO_1-VO_1	###	FO_1-MO_1	**0.7816**	0.1338	###
	FO_2-MO_2-VO_2	**0.9985**	FO_2-MO_2	###	###	###
	FO_3-MO_3-VO_3	**0.9929**	FO_3-MO_3	0.7059	0.2248	0.9658

(*continued*)

Table 2. (*continued*)

Image set	Target set	TRA_VFM	Target pair	NorCroCor	NorCovCor	HOG
Image set 4	FO_1-MO_1-VO_1	**0.9970**	FO_1-MO_1	0.8007	###	0.9943
	FO_2-MO_2-VO_2	**0.9984**	FO_2-MO_2	###	###	###
	FO_3-MO_3-VO_3	**0.9983**	FO_3-MO_3	0.6827	0.0219	0.9747
Image set 5	FO_1-MO_1-VO_1	###	FO_1-MO_1	**0.8305**	0.1895	###
	FO_2-MO_2-VO_2	**0.9923**	FO_2-MO_2	###	###	0.9922
	FO_3-MO_3-VO_3	**0.9681**	FO_3-MO_3	###	0.4501	###
Image set 6	FO_1-MO_1-VO_1	**0.9931**	FO_1-MO_1	###	###	0.9875
	FO_2-MO_2-VO_2	###	FO_2-MO_2	###	###	**0.9646**
	FO_3-MO_3-VO_3	**0.9894**	FO_3-MO_3	###	###	###
	FO_4-MO_4-VO_4	**0.9776**	FO_4-MO_4	0.8960	0.2061	0.9744
Image set 7	FO_1-MO_1-VO_1	**0.9936**	FO_1-MO_1	0.8879	0.4441	0.9863
	FO_2-MO_2-VO_2	**0.9559**	FO_2-MO_2	###	###	###
	FO_3-MO_3-VO_3	0.9744	FO_3-MO_3	###	###	**0.9764**
	FO_4-MO_4-VO_4	0.9695	FO_4-MO_4	0.9119	0.2311	**0.9799**
Image set 8	FO_1-MO_1-VO_1	**0.9939**	FO_1-MO_1	###	###	###
	FO_2-MO_2-VO_2	**0.9803**	FO_2-MO_2	###	###	0.9594
	FO_3-MO_3-VO_3	**0.9899**	FO_3-MO_3	###	###	###
	FO_4-MO_4-VO_4	0.9676	FO_4-MO_4	0.8925	###	**0.9967**
Image set 9	FO_1-MO_1-VO_1	**0.9882**	FO_1-MO_1	###	###	###
	FO_2-MO_2-VO_2	**0.9829**	FO_2-MO_2	###	###	0.9820
	FO_3-MO_3-VO_3	**0.9929**	FO_3-MO_3	0.8877	###	0.9909
	FO_4-MO_4-VO_4	**0.9960**	FO_4-MO_4	0.8702	0.1502	0.9889

It is shown in Table 2 that, in all 31 target sets (target pairs), the TRA_VFM algorithm has the maximum similarity value in 23 target sets; the HOG algorithm has the maximum similarity value in 5 target pairs; the NorCroCor algorithm has the maximum similarity value in 3 target pairs; and the NorCovCor algorithm has no maximum similarity value in any target sets. Overall, the accuracy of target registration of the TRA_VFM algorithm is higher than 0.95. It is demonstrated that the accuracy of target registration of the TRA_VFM algorithm is higher than that of the HOG algorithm, the NorCroCor algorithm and the NorCovCor algorithm.

5.3 Analysis of the Target Registration of VIS and Two Wave Bands Infrared Images

Based on the principle that the same target has the highest similarity among targets in images from different sources, the TRA_VFM algorithm fuses target features in VIS, LWIR and MWIR images by calculating the targets similarity, which improves the registration accuracy of the same target in these images. VIS and LWIR images contain rich edge information and complete geometric shapes information, which can be described by the Hu invariant moment feature. The effect of changes of scale, translation and rotation between images caused by camera motion can be mitigated because

Hu invariant moment feature is invariant to scale, translation and rotation. Hence, the TRA_VFM algorithm calculates the similarity between targets in VIS and LWIR images based on the Hu invariant moment feature and cosine similarity. LWIR and MWIR images contain rich edge information and complete structure information. The improved structural similarity uses images of different scales and their edge images in the Gaussian space to calculate the multi-scale structural similarity, thus it can combine the complete edge structure information of targets in LWIR and MWIR images. Hence, the TRA_VFM algorithm calculates the similarity between targets in LWIR and MWIR images based on the improved structural similarity. The TRA_VFM algorithm calculates the similarity among targets in VIS, LWIR and MWIR images by fusing the two similarities, which realizes the target registration of images of these three sources. Based on the similarity among targets in images from different sources, the TRA_VFM algorithm fuses shape features and structural features of targets in VIS, MWIR and LWIR images. The multi-source information such as edge information, shape information and structural information in these three-source images is fully utilized, which improves the correct rate and the accuracy of target registration. For the proposed TRA_VFM algorithm, the correct rate of target registration is 83.87%, and the accuracy of target registration is higher than 0.95. Both of them are superior to other algorithms. Compared with our algorithm, the investigated references methods only utilize the target information in VIS and LWIR images, lacking of the rich edge information and complete structural information in MWIR images, and comprehensive information of the target is unavailable. Therefore, the target registration result is not as good as our algorithm.

6 Conclusion

In this paper, the shape and structural features of targets in VIS, LWIR and MWIR images are fused based on the similarity among targets in images from different sources, which improves the registration accuracy of the same target in these images. The proposed algorithm and 5 registration algorithms with good performances are tested for target registration performance on 9 sets of VIS, LWIR and MWIR images. Experimental results show that the correct rate and the accuracy of target registration of the proposed algorithm are both superior than those of the other algorithms. In this paper, the target registration of VIS, LWIR and MWIR images is realized, and the correct rate and accuracy of target registration are improved by the registration of multi-source images. In the future, features that better represent target information will be researched in order to realize target registration of images from more than three sources.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61471194 and 61705104), the Fundamental Research Funds for the Central Universities (NJ2020021), the Natural Science Foundation of Jiangsu Province (BK20170804), National Defense Science and Technology Special Innovation Zone Project.

Conflicts of Interest. None.

References

1. Ma, J., Cao, G.Z.: Image matching technology of infrared and visible images based on edge information. Comput. Digit. Eng. **34**, 30–32 (2006)
2. Jiang, J., Zhang, X.S.: Visible and infrared image automatic registration algorithm using mutual information. In: Control and Decision Conference, pp. 1322–1325. IEEE (2010)
3. Mao, L.L., Xu, G.F., Chen, X.B.: Research on multi-source image matching based on fusion of mutual information and gradient. Infrared Technol. **31**(9), 532–536 (2009)
4. Li, Z.: Research on the key technology of heterogeneous image matching, Master Thesis, National University of Defense Technology (September 2011)
5. Bilodeau, G.A., Torabi, A., Charles, P.L., Riahi, D.: Thermal-visible registration of human silhouettes: A similarity measure performance evaluation. Infrared Phys. Technol. **64**, 79–86 (2014)
6. An, M., Jiang, Z., Zhao, D.: High speed robust image registration and localization using optimized algorithm and its performances evaluation. J. Syst. Eng. Electronics **21**(3), 520–526 (2010)
7. Ma, J.Y., Zhao, J., Ma, Y.: Non-rigid visible and infrared face registration via regularized Gaussian fields criterion. Pattern Recogn. **48**(3), 772–784 (2015)
8. Zhuang, Y., Gao, K., Miu, X.: Infrared and visual image registration based on mutual information with a combined particle swarm optimization–Powell search algorithm. Optik-Int. J. Light Electron Optics **127**(1), 188–191 (2016)
9. Dou, J.F., Li, J.X.: Automatic registration of visible and infrared images based on corner and Hu moment invariants. Infrared **32**(7), 23–27 (2011)
10. Wen, X., Xie, T.X., Yan, J.H., Zhang, Y., Huang, W., Chen, X.: Target registration for two wave bands infrared images based on the improved structural similarity. Chinese J. Sci. Instrum. **38**(12), 3112–3120 (2017)
11. Wu, F., Wang, B., Yi, X.: Visible and infrared image registration based on visual salient features. J. Electron. Imaging **24**(5), 053017 (2015)
12. St-Charles, P.L., Bilodeau, G.A., Bergevin, R.: Online multimodal video registration based on shape matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2015)
13. Chen, Y., Zhang, X., Zhang, Y., Maybank, S.J., Fu, Z.: Visible and infrared image registration based on region features and edginess. Mach. Vis. Appl. **29**(1), 113–123 (2017). https://doi.org/10.1007/s00138-017-0879-6
14. Zhang, T., Liu, L., Gao, T.C., Hu, S.: The actuality and progress of whole sky infrared cloud remote sensing techniques. Instrumentation **2**(3), 65–74 (2015)
15. Ma, J.Y., Zhao, J., Ma, Y., Tian, J.W.: Non-rigid visible and infrared face registration via regularized Gaussian fields criterion. Pattern Recogn. **48**(3), 772–784 (2015)
16. Liu, B., Li, W.S.: Indoor location method of fingerprint matching algorithm based on cosine similarity. Bull. Sci. Technol. **33**(3), 198–202 (2017)