An Automatic Registration Algorithm of Infrared and Visible Images Based on Hybrid Image Features

Yahong Zhang [1,2,3,*], Renbo Xia [1,3], Wei Wu [1,2,3]

1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang;
2. University of Chinese Academy of Sciences, Beijing;
3. Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang

ABSTRACT

In this paper, a new method is presented to match a pair of visible and infrared images of same scene based on hybrid visual features including line segments and interest points. First, improved Harris corner extraction method and line segment detector method is used to extract feature points and segments .Then, a novel descriptor integrating the information of line segments and interest points is proposed. Finally, the nearest neighbor algorithm is utilized to match the descriptors, and the RANSAC(Random Sample Consensus) algorithm is employed to rule out the wrong match pairs. The performances are evaluated by extensive experiments on real images. The results show that the proposed algorithm can take advantage of similar structures between the multimodal images to realize automatic matching efficiently.

Keywords: Infrared image; Visual image; hybrid image feature; concentric circle array

1. INTRODUCTION

A fundamental problem in multimodality image integration is to align images of the same scene observed from different positions, different time, and/or in different sensors. Registration between infrared and visible images is an important field of multi-sensor image registrations. The infrared image can record temperature distribution and radiation information of the scene, while the visible image records illumination distribution and reflection information. Consequently, the integration of images from multiple sensors can provide complementary information and therefore increase the accuracy of the overall decision. Additionally, the infrared and visible images are passive imaging and feature good security. These advantages make infrared and visible image registration techniques apply extensively in the fields of military intelligence acquisition, navigation and guidance, remote sensing image fusion, video surveillance and target tracking *etc*.. [9][10]

In accordance with most of the literature, infrared and visible image registration algorithm can be divided into two categories: region-based methods and feature-based methods. Region-based methods directly or indirectly use gray information to calculate space transformation relations between images in a certain similarity measurement. These methods can be further subdivided into gray correlation method, frequency domain method and mutual information method. The first two methods are widely used in single-mode image registration rather than multimodality registration. Mutual Information (MI)[6], which represents statistical correlation of intensity values between images, uses entropy or weighted entropy integrating direction information for judging whether the image is aligned. The method is applicable to various conditions because it does not need too much image pre-processing as well as assumptions on the gray-scale relationship between multi-modal images. The main disadvantage is that MI is intrinsically a global measurement and therefore its local estimation is difficult, which can lead to many false local optima. Moreover, the optimisation of MI is computationally complex and converges slower than simple intensity metrics, such as sum of squared differences (SSD). Feature-based registration method extracts stable geometric elements such as corners, straight, edge structures, contour shape, *etc*, using a certain number of correct association characteristics to deduce space transform relations, or defining the objective function based on feature attributes to identify optimal model parameters. Compared with the region-based method, feature-based methods are advanced in computational efficiency, adaptability and resistance to partial

*E-mail: zhangyahong@sia.cn ; Phone: 024-23970404

International Symposium on Optoelectronic Technology and Application 2014: Image Processing and Pattern Recognition, edited by Gaurav Sharma, Fugen Zhou, Proc. of SPIE Vol. 9301, 93012R · © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2073016

deformation occlusion *etc*, but difficult to extract consistent features. To solve these problems, we make full use of the advantages of various features, and proposed automatic image registration algorithm based on hybrid features including points and lines. Feature-based image registration method comprises three steps: feature detection; feature description and feature matching. We will begin our presentation from the first part of the feature detection.

2. FEATURE DETECTION

In order to improve the robustness of interest point detection, this paper presents the concept of real and virtual corner point. The following are respectively introduced.

2.1 Real corner point detection

Harris[12] proposed a point detection method using the derivative of direction. This method calculates the average gradient square matrix and its eigenvalues of each pixel to determine the corner points. Although this method can effectively detect points of interest, but for the infrared image, the distribution of interest points is often uneven. To solve this problem, this paper proposes an improved Harris corner detection method. An input image is divided into several overlapping blocks with fixed size. In each block, Harris corner extraction method is used to extract interest points. This makes those points that are not globally significant become interest points in a local neighborhood. Using the overlapping method, we can detect the points of interest on the boundary of each block. Furthermore, if a relatively flat region whose maximum value of response function is close to zero, we will discard the detected interest point in this local area. Fig. 1 are results of using two different detection methods. We can see that the use of the sub-area division method can enhance the reasonable distribution of the corners.

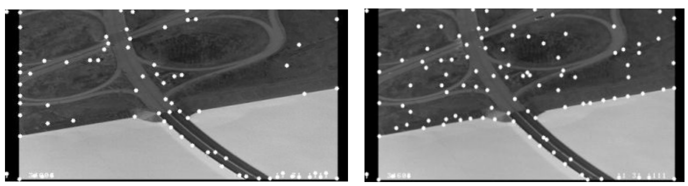
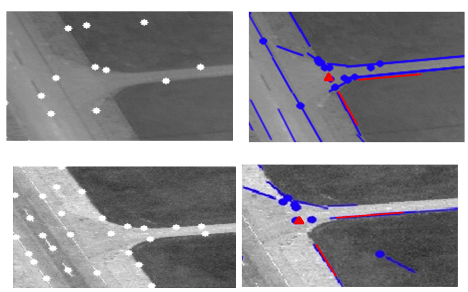


Fig. 1 Different corner extraction results

2.2 Virtual corner point detection



(a) (b)

Fig. 2 (a) Real corner; (b)Virtual corner

Different from the real corner, this paper presents the concept of virtual corner[4]. According to Fig. 2(a), we can see the distribution difference of points of interest detected in infrared and visible images is relatively large, while their edges possess relatively consistent distribution. Thus, this paper presents virtual corner point based on the edge, namely the intersection of the line segments. The definition of virtual corners need to satisfy three constraints: (1) length of segments is not less than the threshold l_{th}. Long segment generally reflect the major structural features of the image, and the long segment's extraction results are relatively robust; (2) the angle between the line need to satisfy $\theta_{t1} < \theta < \theta_{t2}$. Due to the influence of environmental noise, the extracted line may be deviated from the real situation in the position or direction. If segments are close to parallel, even if the deviation is small, the intersection of the segments will be very different with the real situation; (3) the distance between the segments shall not exceed the limit d_{th}. Limit of the distance between the lines can constraint the number of straight line pairs and hence limit the extracted virtual corner point number. If there is no special instruction, the straight line refers to the line segment, namely the straight line between two endpoints. We define the distance between the straight line is the distance between the closest point on the two straight lines. Fig. 2(b) is virtual corner extraction results, it can be seen the extraction corners in the infrared and visible images is highly consistent.

Finally, we merge extraction results of the real corner and virtual corner together to get the final results. As shown in Fig. 3, white marker indicates the results of the improved Harris corner detection method, and the blue marker indicates the virtual corner detection results. It can be seen from the figure, the distribution of the blue markers with higher consistency.

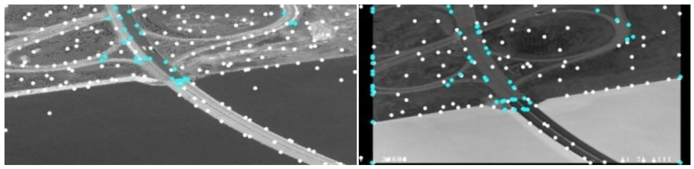


Fig 3 Comprehensive corner detection results

2.3 Line Segment Detection

The LSD algorithm[2] is an excellent line segment detection method based on phase grouping and does not require parameter tuning. In [2], a line segment is defined as an image region whose points share roughly the same gradient orientation. The procedure of the LSD algorithm is given as follows: first, pixels that share the same gradient orientation up to a given tolerance are grouped as a connected region, i.e., the line support region. Then, the line segment that best approximates each line-support region is extracted according to rectangular approximation rules. Finally, each line segment is validated or not based on a contrario model.

3. FEATURE DESCRIPTION

Combination the interest points and line segment detection results, this paper proposes a line context description method. Specifically the line context refers to the distribution of lines in the local neighborhood centered at the interest point. It includes single distribution and the family distribution of the line segments. Single line description is to quantify a line segment with a vector or scale, while the family line description describes the positional relationship between segments.

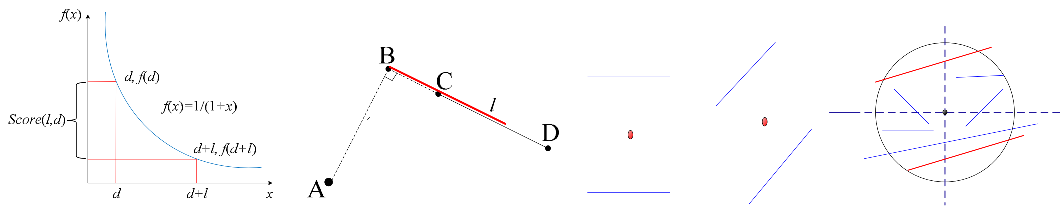
3.1 The description of the single line distribution

We use three properties of a line, including the length, the direction, and the distance to the center point, to describe the distribution on a single line[11]. Based on these three properties we construct a score function to describe the distribution, as shown in (1),

$$Score(l,d) = \frac{1}{1+d} - \frac{1}{1+d+l}, \tag{1}$$

Based on the distribution of the line relative to the interest point, the design of score function must follow two principles: the longer the line segment, the higher the score; the shorter the distance to the interest point, the higher the score. As can be seen from the Fig 4(a), with the increase of l, $f(d+l)$ move down and the score increased; as d decrease, $f(d)$ and $f(d+l)$ move up simultaneously, but $f(d)$ move faster than $f(d+l)$ and the score increased; which conform to the two design principle.

There is an important variable d in the score function, where d is Manhattan distance (AB+BC) of the point A to segment CD. We use Manhattan distance (AB+BC) instead of Euclidean distance AB here because if we use AB, the red line in Fig. 4(b) which has an equal length with, they will have equal score. Manhattan distance can effectively distinguish the two different line distribution. Finally, the score of the line is divided to x and y direction to indicate direction information of the line.



(a) (b) (c) (d)

Fig. 4 (a) Score of a line segment to interest point; (b) Illustration of Manhattan distance; (c) Different line segment with same score; (d) Four quadrant characteristic description method

3.2 Description of the family line distribution

There are conventionally numerous line segments in the local neighborhood. After single line distribution is described, we consider the spatial relationship between line segments. As shown in Figure 4(c), the two lines with equal length, equal distance and equal angle corresponding to the interest point will obtain a same score, which means different segments distribution corresponding to the same score.

In order to prevent the above condition, this paper puts forward four quadrant characteristic description method, as shown in Fig. 4(d), each line segment in this neighborhood is partitioned into several sub-segments in different quadrants and sub-segments are described respectively to distinguish spatial position relationship between the line segments. As in the Fig 4(d) this method can effectively differentiate spatial distribution of lines labeled in red color. Then score of each segment is added to indicate the distribution of the line in each quadrant. Finally, the score is transformed into a feature descriptor to indicate the distribution of segment in this region.

3.3 Robust Features Description

If constructing descriptor with line segments only in the local neighborhood, we can only get low dimensional descriptors. In order to improve the robustness of descriptor, we propose the concept of concentric array. In SIFT[1], centering on the interest point, we selected 16 rectangular area and obtain gradient direction histogram in each sub-domain to get a higher dimensional descriptor. In this paper, learn from this idea, we construct several concentric circles with different radius, and select uniform sampling points on the concentric circles. Then each sample point will be described in the same way as the interest point. We will put sampling point and interest points together to obtain a higher dimensional descriptor. The descriptor can describe all line segments distribution around the interest points and sample points. The circular array has more advantages compared to rectangular array in SIFT which is preferred to achieve rotation and scale invariance by rotating samples on the circle and adjusting circle radius. Finally descriptor is normalized to resist the influence of illumination change.

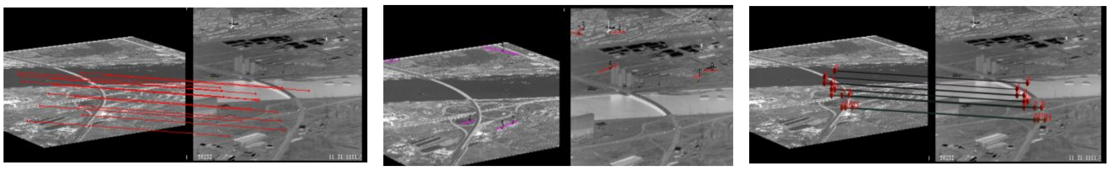
4. FEATURE MATCHING

This process finds the matches between keypoints using the Euclidean distance between the corresponding descriptors. The nearest neighbor algorithm is utilized to match the descriptors. After getting the initial matching points, there are still many wrong matching points. The RANSAC algorithm[7] is employed to rule out the wrong match point pairs.

5. EXPERIMENTS

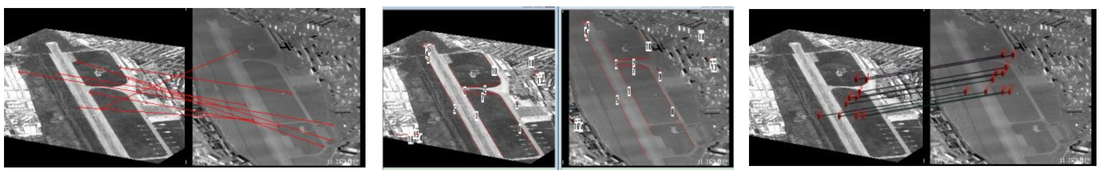
Using the above matching algorithm, visible and infrared images are matched. The algorithm performance is analyzed from the time complexity and robustness which is compared with two classical multimode image matching algorithm respectively. One is image matching method based on the local self-similarity[3] and the other is wide baseline image matching method based on line feature[5]. The comparison results of three methods are illustrated in the Fig. 5 and Fig. 6. According to the comparison result , due to the line feature is not obvious or the segments is too short, the matching method based on line feature is not good. In the second comparison results, LSS method cannot identify structural areas with similar texture, and in contrast, the method in this paper has achieved good results.

To sum up, LSS can't recognize some structure information of similar texture and is difficult to realize the rotation and scale invariance. Line Signature is scale and rotation invariance, but it would fail if the line segment is not obvious or too short. Furthermore, inaccurate endpoint location will cause errors because this method is dependent on the endpoint location. The proposed method in this paper can be applicable to the general situation and has higher robustness, and easy to realize the rotation and scale invariance. In the aspect of time-consuming, the method of this paper is slightly better than other two methods, as shown in table 1, but still need further improve.



(a) LSS (b) Line signature (c) Method in this paper

Fig. 5 The first set comparative results



(a) LSS (b) Line signature (c) Method in this paper

Fig. 6 The second set comparative results

Table 1 Comparison of time performance

Methods	**LSS**	**Line Signature**	**Method in this paper**
Average time	20.24 s	12.10 s	10.38 s

6. CONCLUSION

In this paper, a new method is presented to match a pair of visible and infrared images of same scene based on hybrid visual features including line segments and interest points. The results show that the proposed algorithm can take advantage of similar structures between the multimodal images to realize automatic matching efficiently. The comparisons with other methods indicate that our approach is suitable for general situations and has a high robustness.

Acknowledgements

The Work is fund by Chinese Natural and Science Fund grant:.

References

[1] Lowe, D. G., "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60(2), 91-110(2004).
[2] Von Gioi, R. G., Jakubowicz, J., Morel, J. M. and Randall, G., "LSD: A fast line segment detector with a false detection control,". Pattern Analysis and Machine Intelligence, 32(4), 722-732(2010).
[3] Shechtman, E., and Irani, M., "Matching local self-similarities across images and videos," CVPR, 1-8 (2007).
[4] Gao, F., Wen, G. J. and Lv, J. J., "An optimal algorithm for IR/ Visual image registration based on main -line -pairs," Chinese Journal of Computers, 30(6), 1015-1021(2007).
[5] Wang, L., Neumann, U. and You, S., "Wide-baseline image matching using line signatures," ICCV, 1311-1318(2009).
[6] Viola, P. and Wells III, W. M., "Alignment by maximization of mutual information," International Journal of Computer Vision, 24(2), 137-154(1997).
[7] Wang, H. and Suter, D., "Robust adaptive-scale parametric model estimation for computer vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(11), 1459-1474(2004).

[8] Han, J., Pauwels, E. J. and De Zeeuw, P., "Visible and infrared image registration in man-made environments employing hybrid visual features," Pattern Recognition Letters, 34(1), 42-51(2013).
[9] SU, K. F., DENG, Z. D. and HUANG, Z., "A novel localization approach for autonomous vehicles based on map matching with curvature features," ROBOT,34(4), 440-448(2012).
[10] ZOU, Y. H., CHEN, W. H., WU, X. M. and LIU, Z., "A Stereo Matching Algorithm under Outdoor Unideal Illumination Condition," ROBOT, 34(3), 344-353(2012)
[11] Pang, G. and Neumann, U., "The Gixel array descriptor (GAD) for multimodal image matching," WACV 497-504(2013).
[12] Derpanis, K. G., "The harris corner detector," York University(2004).