AN IMPROVED IMAGE MATCHING METHOD BASED ON SURF ALGORITHM

S. J. Chen[1,*], S. Z. Zheng[2], Z. G. Xu[1], C. C. Guo[2], X. L. Ma[3]

[1] School of Resource Engineering, Longyan University, Longyan, China - 13806984291@139.com, xzg_xzg1982@163.com
[2] College of Surveying and Geo-Informatics, Tongji University, Shanghai, China - (zhengshouzhu0, candyguo_fly)@163.com
[3] Chinese Academy of Surveying and Mapping, Beijing, China -1510912@tongji.edu.cn

Commission III, WG III/6

KEY WORDS: Image matching, Information entropy, SURF, Delaunay Triangulation

ABSTRACT:

Many state-of-the-art image matching methods, based on the feature matching, have been widely studied in the remote sensing field. These methods of feature matching which get highly operating efficiency, have a disadvantage of low accuracy and robustness. This paper proposes an improved image matching method which based on the SURF algorithm. The proposed method introduces color invariant transformation, information entropy theory and a series of constraint conditions to increase feature points detection and matching accuracy. First, the model of color invariant transformation is introduced for two matching images aiming at obtaining more color information during the matching process and information entropy theory is used to obtain the most information of two matching images. Then SURF algorithm is applied to detect and describe points from the images. Finally, constraint conditions which including Delaunay triangulation construction, similarity function and projective invariant are employed to eliminate the mismatches so as to improve matching precision. The proposed method has been validated on the remote sensing images and the result benefits from its high precision and robustness.

1. INTRODUCTION

Image matching is a process of a geometrically matching up two images of a same scene, which widely used in various applications (Zitova and Flusser, 2003; Arévalo and González, 2008), including image mosaicing, change detection, 3D construction, etc. Generally speaking, image matching algorithms can be classified to feature-based and intensity-based methods (Zitova and Flusser, 2003; Wu et al., 2015). Compared with intensity-based algorithms, feature-based algorithms not only have advantages of illumination changes and complex distortion, but also take less time consuming due to the whole image information replaced by a finite number of points. So feature-based algorithms have been a research hotspot of image matching. Feature-based algorithms firstly extract feature points from two images, and the matching method is performed to generate the matching result from the feature points pair.

In recent years, feature-based algorithms have been extensively studied. Lowe proposed scale invariant features transform (SIFT) algorithm (Lowe, 2004) to solve some problems such as illumination, speckle, rotation, scale, translation, etc. However, the algorithm is not fit for dealing with a large number of images because of SIFT itself complexity. Then many researchers have already been attracted to improve SIFT (Mikolajczyk and Schmid, 2005; Juan and Gwun, 2009; Teke et al., 2011). Speeded Up Robust Features (SURF) algorithm (Bay and Van, 2006) is the improved version of SIFT, having an advantage of time efficiency which is attributed to introduce integral image and box filter to reduce time-consuming. But it has a disadvantage of lower accuracy. Lee et al. (Lee et al., 2010) presents a Coarse-to-Fine approach of image matching based on Haar Wavelet Transform and SURF algorithm. The Coarse-to-Fine strategy from Harris operator to normalized cross-correlation and RANdom SAmple Consensus (RANSAC) algorithm are used to achieve the fine points. Bouchiha et al.

developed a matching method which separated the detector from the descriptor and proposed an extension to the SURF descriptor. Zheng et al. (Zheng et al., 2015) proposed an improved SURF method by combining color invariant model and a series of constraint conditions (CC_SURF), this method has a good matching rate and feature points are well-distributed. Anzid et al. (Anzid et al., 2017) proposed an improvement of the SURF algorithm and it could automatically remove the outliers by means of both Distance and Orientation filtering strategy (DO_SURF). However, in the aforementioned methods the number of the detected feature points is not at a high level owing to the loss of color information caused by color images often converting to grayscale images in the matching process.

Therefore, this paper proposes an improved SURF mean which introduces color invariant transformation, information entropy and constraint conditions aiming at improving a number of feature points and matching precision. First, orthogonal color transformation model is introduced to generate the orthogonal color space of two images from RGB color space of both images, and the information entropy is used to selected the most rich color channel as the subsequent matching image. Then SURF algorithm is applied to extract and describe feature points from the image, and constraint conditions which including constructing Delaunay triangulation, triangle similarity function and projective invariant are employed to eliminate the mismatch match and increase image match precision and robustness.

The rest of this paper is organized as follows: Section 2 describes the proposed method. In Section 3, experimental results obtained on remote sensing images by the proposed method is compared with other related SURF algorithms. Section 4, the conclusion is drawn.

2. THE PROPOSED METHOD

2.1 Overview of the proposed method

In this paper, the overall flowchart of the proposed method is shown in Fig. 1. It can be divided into 3 main steps: (1) Color transformation and Information entropy are introduced for obtaining the more information data of two matching images; (2) SURF algorithm are used to detect and extract the feature points; (3) methods of mismatch points removal (Li and Zhang, 2009; Zheng et al., 2015) which including Delaunay triangulation, triangle similarity function and projective invariant, are applied to eliminate the mismatch feature points.
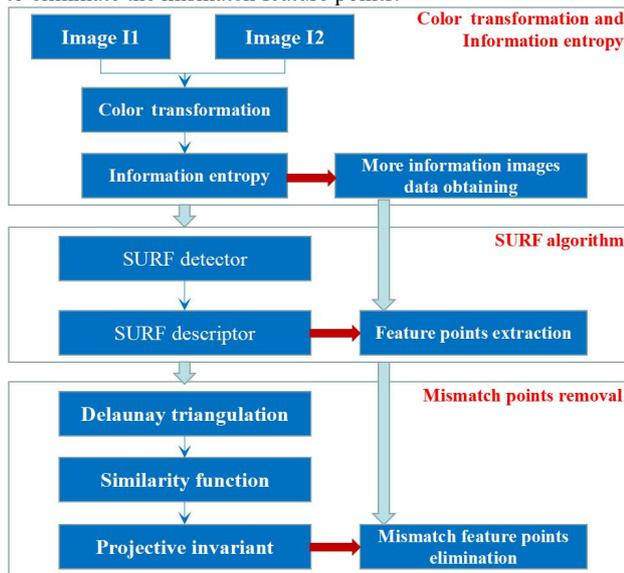

Fig. 1. The overall flowchart of the proposed method

2.2 Color transformation and Information entropy

Due to the RGB color space of images without invariant space, we transform RGB color space into orthogonal color space with color invariant characteristic. So each color channel of image which including R, G and B band, is done by calculated the mean and the standard deviation of pixel value. The orthogonal color space calculation expression can be defined as:

$$\begin{bmatrix} R^0 \\ G^0 \\ B^0 \end{bmatrix} = \begin{bmatrix} (R - \mu_R)/\delta_R \\ (G - \mu_G)/\delta_G \\ (B - \mu_B)/\delta_B \end{bmatrix} \tag{1}$$

where μ_R, μ_G and μ_B are the mean of pixel value in each channel, respectively. δ_R, δ_G and δ_B are the standard deviation value of each band respectively. R, G and B respectively represent the pixel value of every band in original image. R^0, G^0 and B^0 are the pixel value of every band via orthogonal transformation.

Information entropy is calculated to select the most abundant information for each channel. Information entropy is a measure of image information. The greater the information entropy of an image, the more the information content has. Therefore information entropy is introduced to calculate the information of each channel and select the best channel as import information data to be processed. More feature points would be detected by SURF detector in the both images. Information entropy expression can be described as:

$$E(j)_k = -\sum_{j=\min}^{\max} \rho(j) \ln \rho(j), k = R^0, G^0, B^0 \tag{2}$$

where k represents the three corresponding color channels; j represents pixel value level of color channels; $\rho(j)$ represents occurrence probability of j-level; min and max are the pixel value of minimum and maximum, respectively. $E(j)$ represents information entropy of color channels.

As the above mentioned by equation 1 and 2, the most entropy of color channels is used for the feature detection and extraction of the next step.

2.3 SURF detector and descriptor

2.3.1 SURF detector: For the two matching images, we calculate the information entropy for all the channels and select the channel with the maximum entropy as input data. SURF detector is used to detect the feature points of two image. Integral image is set up and multi-scale space is built by box filter. Hessian matrix is applied as it has good performance and accuracy. The Hessian matrix can be defined as:

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \tag{3}$$

where σ is a scale. $L_{xx}(x, \sigma)$ is the convolution result of the second order derivative of Gaussian filter, and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$.

In order to improve the calculation efficiency, we construct a Fast-Hessian matrix in SURF detector. We equate the initial 9×9 box filter with σ=1.2 of second order derivative of Gaussian filter and D_{xx} is a result which received by convolution operation of box filter template and input image, represents L_{xx}, similarity for D_{xy} and D_{yy}. The Fast-Hessian matrix can be expressed as follows:

$$\det(Hessian) = D_{xx}D_{yy} - (0.9D_{xy})^2 \tag{4}$$

Finally, the threshold is appropriately selected and a Non-maximum suppression in a 3×3×3 neighborhood of each point is applied to detect extremum points. The point is regarded as a candidate only if its response value is larger or smaller than all the 26 neighbors at the current and adjacent scales. Then, the steady feature point location and its scale are obtained by interpolating in scale space and image space.

2.3.2 SURF descriptor: To realize the rotation invariant of feature point, the dominant orientation for each feature point should be determined. First, Harr-wavelet responses are calculated for the pixels within a circular neighbourhood of radius 6s around the feature points, where refers to the scale of point detected. Then weighted Gaussian function is performed for the feature points, and we give greater weights only if the feature point is close to the circle center and contributing a lot in orientation. The dominant orientation is estimated by calculating the sum of all responses within a sliding orientation window covering an angle of 60°. The x- and y- responses within the window are summed up to produce a new vector. So the longest vector's orientation is selected as the feature point dominant orientation via scanning the entire circular neighbourhood.

After generating feature descriptor, the rectangular region with the side length 20s×20s is selected, and the dominant orientation of the area are rotated to the dominant orientation of the feature points. The rectangular region is divided into 16 smaller 4×4 square sub-regions and 5×5 sampling points are selected in each sub-region to calculate the corresponding Harr-wavelet responses d_x and d_y. Furthermore, summate Harr-wavelet response values and their absolute values of the 4 sub-regions separately. So we get a four-dimension descriptor vector $v=(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$ for each sub-regions and form a 64 (4×4×4)-dimension feature vector.

## 2.4 Mismatch points removal

SURF detector and descriptor only detects and extracts feature points, but it can hardly eliminate a lot of mismatch points. For this reason, this paper introduces the Delaunay triangulation, triangle similarity function and projective invariant under the matching process.

2.4.1 Delaunay triangulation construction: In this paper, incremental insertion algorithm (Lawson, 1977) which has some advantages of simple thought, easy implementation and high efficiency, is used to build triangulation network for all feature points in both images. Then LOP (Local Optimization Procedure) is used to optimize the quality and performance of Delaunay triangulation. The procedure of triangulation building and LOP optimizing are defined as follows:

(1) Incremental insertion algorithm: As shown in Fig. 2, the basic steps of building triangulation which use the incremental insertion algorithm, is described in Table 1:

Step1	build a large triangle as the initial triangle which contains all of the points.
Step2	freely select one from the points as interpolation point in the large triangle.
Step3	search the triangle which includes this point, then link this point with other three point of the triangle to generate three new small triangle.
Step4	call the LOP optimization approach to update all the triangle generated by Step2.
Step5	repeat Step2 to Step4 until all other points are to process.
Step6	delete the triangles which contains initial triangle vertices to generate Delaunay triangulation network (Fig. 2(f)).

Table 1 Basic steps of incremental insertion algorithm

(a) build an initial triangle (b) generate an initial triangulation network

(c) insert the second feature point (d) insert the third feature point
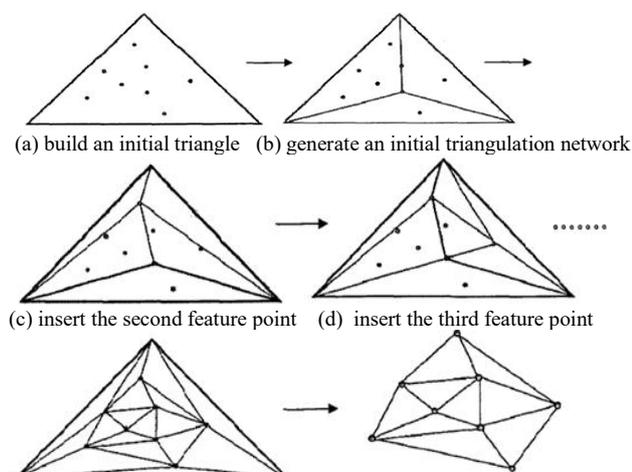
(e) insert the n-th point by iteration (f) Delaunay triangulation network
Fig. 2. Incremental insertion algorithm schematic diagram

(2) LOP (Local Optimization Procedure): LOP is based on the nature of maximum-minimum angle, which means that the minimum of the six angles won't be increased by exchanging two diagonals of a convex quadrilateral. So we can usually use the related properties of angle of circumference to judge whether the minimum of the six angles changes or not. As shown in Fig. 3(a), when the point D locates at the circumscribed circle of the triangle ABC, it can meet the condition of $\angle A + \angle D = \pi$ and $\sin(\angle A + \angle D) = 0$. So we could not exchange two diagonals of a convex quadrilateral. In Fig. 3(b), when the point D lies outside of the circumscribed circle of the triangle ABC, two diagonals would not be also exchanged due to meet the condition of $\angle A + \angle D < \pi$ and $\sin(\angle A + \angle D) > 0$. When the point D locates in the circumscribed circle as shown in Fig. 3(c), we can exchange two diagonals to meet this nature of minimum angle.

(a) Point D locates at the circle (b) Point D locates outside of the circle
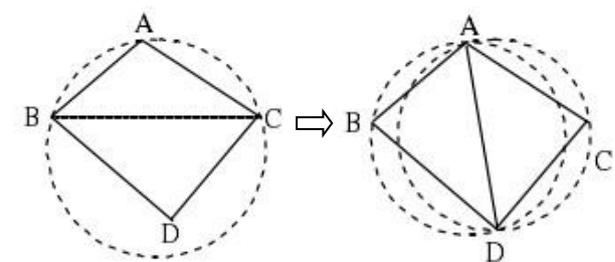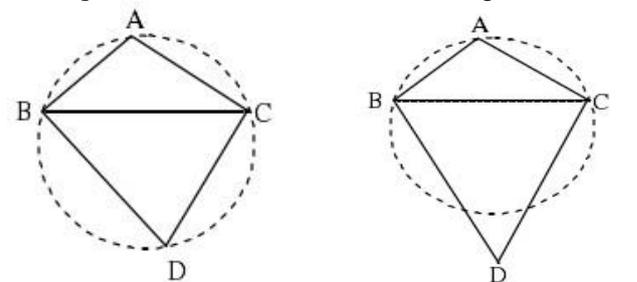
(c) Point D locates in the circle
Fig. 3. LOP judgment of three cases

2.4.2 Triangle similarity function: Triangles generated by the incremental insertion algorithm in triangulation of both images, have the great geometric similarity. Then we can eliminate the mismatched points via this similarity of triangle.

For the two similar triangles of ABC and A`B`C` (A corresponds to A`, similarly for B and C), the similarity I_a of $\angle A$ (its value is a) and $\angle A`$ (its value is x) is defined as:

$$I_a = \cos^3(\frac{\pi}{2}(1 - d(x))) \tag{7}$$

where $d(x)$ equals to $exp\{-(x-a)^2/2\sigma^2\}$, σ equals to $a/6$.

For a pair of triangles, the similarity I_i of other two angles can also be calculated by equation (7). So the similarity of both triangles can be indicated as equation (8).

$$I = (I_a + I_b + I_c)/3 \tag{8}$$

In this paper, we search and select the triangles which are greater than 0.75.

2.4.3 Projective invariant: According to the projection relation of the two images, projective invariant can be used to judge whether points remain the same nature and quantity after the projective transformation. Based on the above process, projective invariant process can make the two images to obtain fine matching. In this paper, cross-ratio is used to analyze a pair of triangles from two images. As stated above, the cross-ratio of straight line to point A and the corresponding A' is taken as an example, as shown in Fig. 4.

(a) Point A triangle and its collinear adjacent triangle
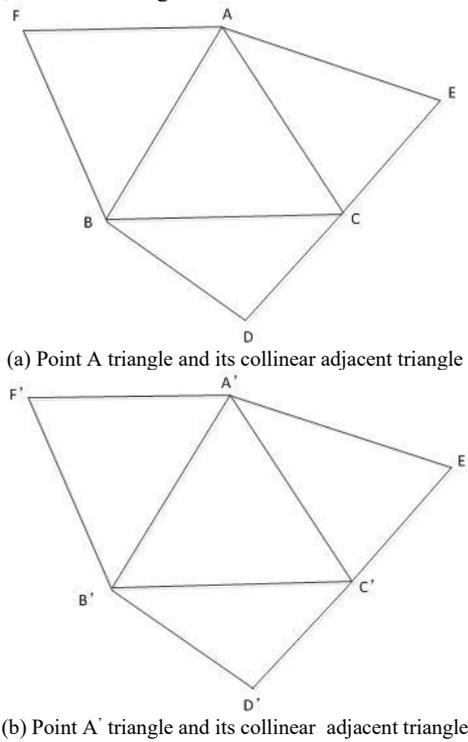
(b) Point A' triangle and its collinear adjacent triangle
Fig. 4. The triangle and the collinear adjacent triangle

From Fig. 4(a) and Fig. 4(b), the cross-ratio of point A represents $I_A=(\sin\angle FAC*\sin\angle BAE)/(\sin\angle FAE*\sin\angle BAC)$, and the cross-ratio of point A' expresses $I_{A'}=(\sin\angle F'A'C'*\sin\angle B'A'E')/(\sin\angle F'A'E'*\sin\angle B'A'C')$. These straight lines which consist triangles are regarded as the right match, only if I_A equals to $I_{A'}$. Finally, the feature points which compose of these straight lines, are accurately extracted from both images.

3. EXPERIMENTAL RESULTS

In order to verify the reliability and advantage of the proposed method, three reference methods (SURF+RANSAC, CC_SURF, DO_SURF) and the proposed approach are compared to analyze their performance and the obtained results. Then evaluation criterion which consists of correct matches number (N) (Li et al., 2015; Ma et al., 2017) and root mean square error (Gong et al., 2014; Kupfer et al., 2015; Ma et al., 2017), is used to verify the accuracy and robustness of the proposed method. In the paper, we selected the two images with different perspectives that both images as shown Fig. 5 (Zheng et al., 2015). These four methods stated above are used to perform and operate matching process of the both image, and the corresponding results of feature points detection, feature points matching and matching results are shown in Fig. 6, Fig. 7 and Fig. 8, respectively. The

corresponding value of N and RMSE are shown in Table 2, where Fig. 2 only lists the results of feature points detection with Image I22.
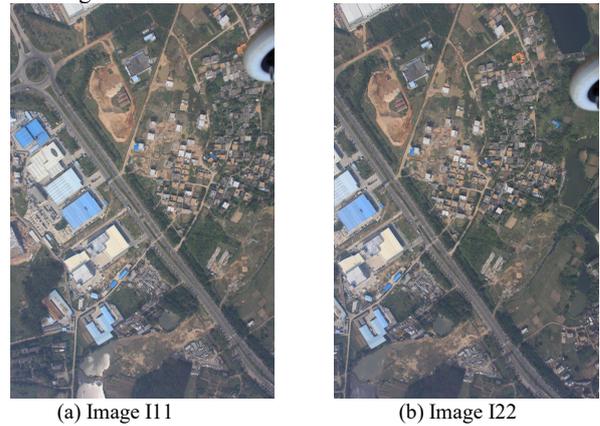

(a) Image I11   (b) Image I22
Fig. 5. Original images (Zheng et al., 2015)

(a) SURF+RANSAC   (b) CC_SURF
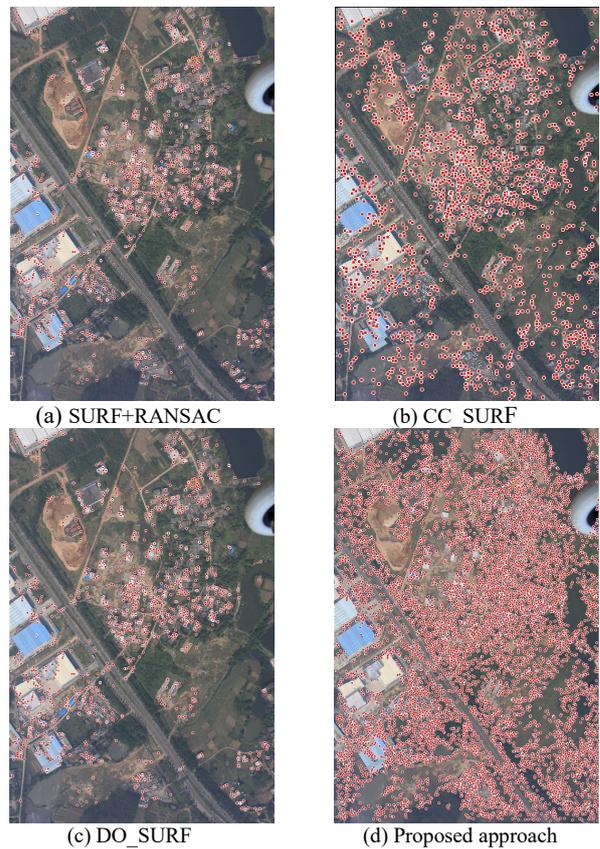
(c) DO_SURF   (d) Proposed approach
Fig. 6. Feature points detection of three methods

(a) SURF+RANSAC   (b) CC_SURF

(c) DO_SURF   (d) Proposed approach
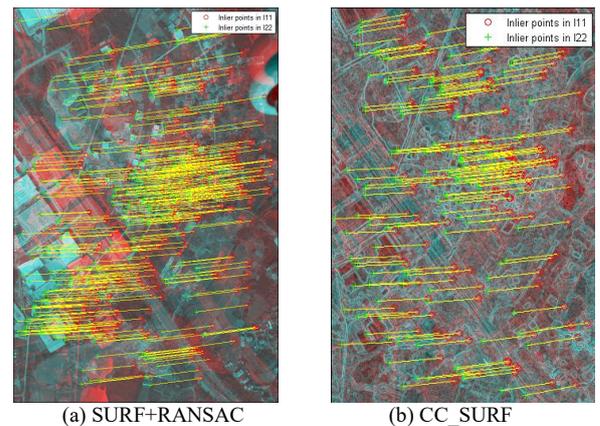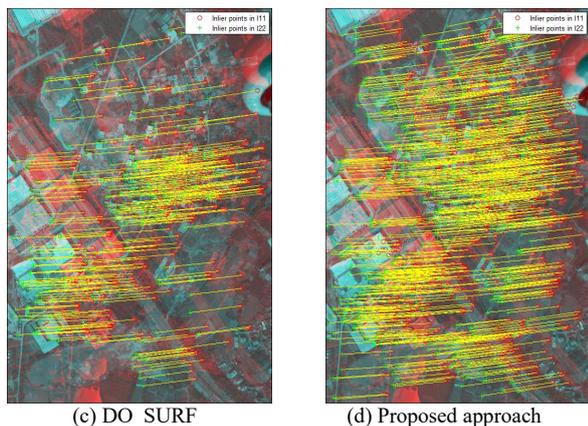
Fig. 7. Feature matching of three methods

(a) SURF+RANSAC   (b) CC_SURF

(c) DO_SURF   (d) Proposed approach
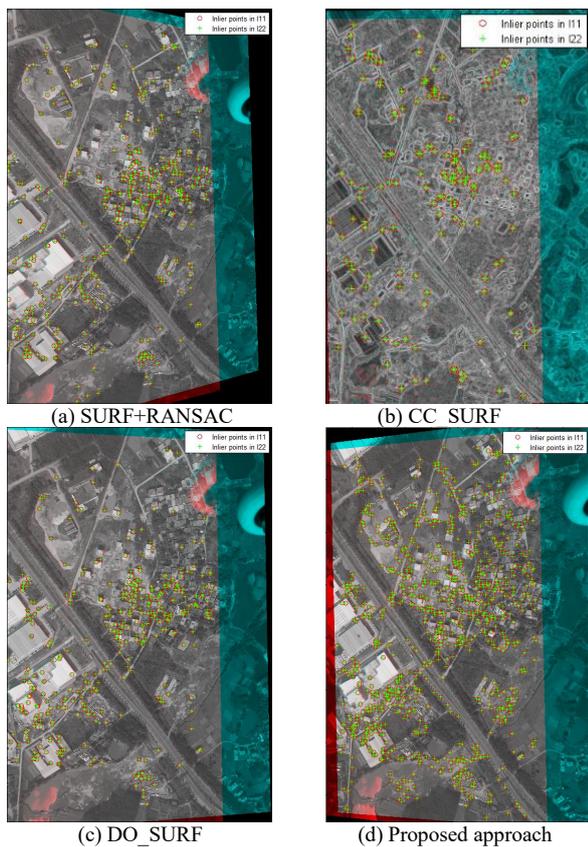
Fig. 8. Matching results of three methods

Methods	N	RMSE (pixel)
SURF+RANSAC	424	0.9253
CC_SURF	166	0.8647
DO_SURF	323	08983
Proposed method	893	0.7507

Table 2 Comparison of N and RMSE

From Fig. 6, we can see that the proposed method detected about ten times feature points more than other three methods due to the more rich color information obtained by color invariant transformation and information entropy. Fig. 7 and Fig. 8 show that all four methods greatly removed many outliers and remain correct matches point pairs. However, the number of correct matches obtained by the proposed method, are more than those by three comparison methods, and the evenness of match points obtained by the proposed approach, are better than the others. In addition, as shown in Table 2, the RMSE value of

SURF+RANSAC, CC_SURF and DO_SURF are larger than the proposed approach. On the contrary, the Correct Matches Number (N) of SURF+RANSAC, CC_SURF and DO_SURF are less than the proposed approach. It can be concluded that the proposed approach outperforms the three comparison methods in terms of the correct matches number and matching accuracy.

4. CONCLUSION

In the paper, we propose a method based on SURF to improve detecting and matching performance. The proposed method introduces color invariant transformation and information entropy to greatly maintain color information and detect more feature points. Then a series of constraint conditions including Delaunay triangulation, triangle similarity function and projective invariant, are used to filter out the mismatch feature points, and are to ensure a high correct matching accuracy of matching results. The above experiments also showed that the robustness and precision of the proposed approach are superior to the others three methods.

REFERENCES

Anzid, H., Le Goic, G., Bekkari, A., Mansouri, A., and Mammass, D., 2017. An automatic filtering algorithm for SURF-based registration of remote sensing images. In: 3rd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP2017), Morroco, Venezuela, pp. 1-7.

Arévalo, V., and González, J., 2008. An experimental evaluation of non-rigid registration techniques on Quickbird satellite imagery. *International Journal of Remote Sensing*, 29(2), pp. 513-527.

Bay, H., Tuytelaars, T. and Van G. L., 2006. Surf: Speeded up robust features. In European conference on computer vision Springer, Berlin, Heidelberg, pp. 404-417.

Besl, P. J., and McKay, N. D., 1992. Method for registration of 3-D shapes. In: International Society for Optics and Photonics, Sensor Fusion IV: Control Paradigms and Data Structures, Vol. 1611, pp. 586-607.

Gong, M., Zhao, S., Jiao, L., Tian, D., and Wang, S., 2014. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7), pp. 4328-4338.

Juan, L., and Gwun, O., 2009. A comparison of sift, pca-sift and surf. *International Journal of Image Processing* (IJIP), 3(4), pp. 143-152.

Kupfer, B., Netanyahu, N. S., and Shimshoni, I., 2015. An efficient SIFT-based mode-seeking algorithm for sub-pixel

registration of remotely sensed images. *IEEE Geoscience and Remote Sensing Letters*, 12(2), pp. 379-383.

Lawson, C. L., 1977. Software for C^1 surface interpolation. In Mathematical software, pp. 161-194.

Lee, S. R., 2010. A Coarse-to-Fine Approach for Remote-Sensing Image Registration Based on a Local Method. *International Journal on Smart Sensing & Intelligent Systems*, 3(4), pp. 690-702.

Li, Q., Qi, S., Shen, Y., Ni, D., Zhang, H., and Wang, T., 2015. Multispectral image alignment with nonlinear scale-invariant keypoint and enhanced local feature matrix. *IEEE Geoscience and Remote Sensing Letters*, 12(7), pp. 1551-1555.

Li S. K. and Zhang P., 2009. Image matching algorithm based on delaunay triangulation and projective invariant. *Microcomputer Information*, 25(28), pp. 115-117.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), pp. 91-110.

Ma, W., Wen, Z., Wu, Y., Jiao, L., Gong, M., Zheng, Y., and Liu, L., 2017. Remote sensing image registration with modified sift and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, 14(1), pp. 3-7.

Mikolajczyk, K., and Schmid, C., 2005. A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, 27(10), pp. 1615-1630.

Teke, M., Vural, M. F., Temizel, A., and Yardimci, Y., 2011. High-resolution multispectral satellite image matching using scale invariant feature transform and speeded up robust features. *Journal of Applied Remote Sensing*, 5(1), pp. 053553:1-9.

Wu, Y., Ma, W., Gong, M., Su, L., and Jiao, L., 2015. A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience and Remote Sensing Letters*, 12(1), pp. 43-47.

Zheng, S. Z., Cheng, P. G., Chen, X. Y., and Chen, S. J., 2015. An approach of image matching combining with surf and delaunay triangulation. *Science of Surveying & Mapping*. 40(2), pp. 155-159.

Zitova, B., and Flusser, J., 2003. Image registration methods: a survey. *Image and vision computing*, 21(11), pp. 977-1000.