International Journal of Engineering

Journal Homepage: www.ije.ir

DPML-Risk: An Efficient Algorithm for Image Registration

S. Kazemi*, M. R. Ahmadzadeh

Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

PAPER INFO

Paper history:
Received 16 March 2018
Received in revised form 30 August 2018
Accepted 26 October 2018

Keywords:
Feature Detection
Binary Descriptor
Image Registration
Scale Invariant Feature Transform

ABSTRACT

Targets and objects registration and tracking in a sequence of images play an important role in various areas. One of the methods in image registration is feature-based algorithm which is accomplished in two steps. The first step includes finding features of sensed and reference images. In this step, a scale space is used to reduce the sensitivity of detected features to the scale changes. Afterward, we attribute feature points that obtained in the first step, descriptions using brightness value around the feature points. In this paper, a new algorithm is proposed based on Binary Robust Invariant Scalable Keypoints (BRISK) and Scale Invariant Feature Transform (SIFT) algorithms. The proposed algorithm uses the directional pattern to describe the edges which are around the keypoints. This pattern is perpendicular to the direction of keypoints which shows the direction of the edge and provides more useful information regarding brightness around the feature point to make descriptor vector. Furthermore, in the proposed algorithm, the output vector consists of multilevel values instead of binary values which means further useful information is involved in the descriptor vector. Also, levels of output vectors can be adjusted using a single parameter so that the processor with low computing ability can tune the output to a binary vector. Experimental results show that the proposed algorithm is more robust than the BRISK algorithm and the efficiency of the algorithm is about the same as BRISK algorithm.

doi: *10.5829/ije.2018.31.11b.09*

1. INTRODUCTION

Image registration can be defined as overlaying two or more images taken at different times, by different sensors or sources or from different points of view. Image registration, also known as image fusion, matching or warping, can be defined as the process of aligning two or more images [1]. Due to the variety of different types of images, designing a unified and general purpose image registration approach is very difficult.

The method which is designed for an image registration application depends on the geometric transformation between the images, the amount of noise damage, accuracy, and application. Image registration algorithms are usually implemented in five steps [2]; preprocessing, feature detection, feature description, feature matching, image resampling, and transformation.

In the first step, some algorithms like smoothing, de-blurring, edge detection and segmentation are used to prepare images for registration. Feature detection is the next step. Feature in an image is a region of the image which is conceptually interesting, or in other words the abstract region of an image is a feature [3]. These features may be a set of corners, interest points, contours, edges, regions or larger features like blobs [4]. Furthermore, feature detection algorithms are preferred to have properties such as robustness, repeatability, accuracy, generality, efficiency, and quantity [5].

Then, a descriptor vector for each detected feature is provided. The descriptor vector is desired to be unique and be robust to some common changes in images. In the next step of the registration algorithms, features are matched using descriptor vector. If the number of matched features in the reference and sensed images are more than a threshold, then a proper transformation is

*Corresponding Author Email: Kazemi_Saeed@Hotmail.com (S. Kazemi)

applied to warp the sensed image to the reference image.

In this study, a new algorithm called Directional Pattern Multi-Level Robust Invariant Scalable Keypoints (DPML-RISK) is presented for image registration. The algorithm consists of two steps, keypoints detection and description. In the keypoints detection step, the Difference of Gaussians (DOG) algorithm is used, but generally, any detector which is insensitive to the scale changes can be used.

The proposed algorithm uses a directional pattern to describe the feature. First, the gradient of the keypoints which shows maximum brightness variations is obtained. Then the directional pattern is rotated to align with the gradient of the keypoints and finally, a vector is made by comparing the brightness of pixels with each other. This new idea provides more useful information regarding the brightness around the feature to make a descriptor vector. The reason is that the brightness around the feature indicates an edge, and the gradient of the keypoints shows the direction of the edge. So using directional sampling pattern tries to encode the edges around the keypoints. In other words, the presented algorithm describes edges of the keypoints instead of the keypoints. Also to improve discriminatory power of the algorithm, multi-level descriptor vector is used.

In section 2, the related works are reviewed. In the third section implementation of the proposed algorithm is explained in detail and the experimental results are examined in the last section.

2. RELATED WORKS

One of the well-known local feature detectors used in image registration algorithms is SIFT (Scale Invariant Feature Transform), which was presented by David Lowe [6]. This algorithm uses DOG function for feature detection. DOG function is a good approximation for gradient with low computational time [7]. The descriptor of this algorithm is based on the histogram of the gradient of pixels around the feature.

Bay et al. [8] introduced the Speeded-Up Robust Features (SURF) detector-descriptor. In this algorithm, the Haar wavelet is used to describe the keypoint. These two methods use Histogram of Gradient (HOG) [9]. Binary Robust Invariant Scalable Keypoints (BRISK) is another binary descriptor, which was introduced by Leutenegger et al. [10]. The algorithm uses the scale-space and the Adaptive and Generic Accelerated Segment Test (AGAST) corner detector [11] to find the features robust to the scale changes. For this purpose, a pyramid of images is organized, and then the AGAST corner detector with the same thresholds is applied on all layers of the pyramid. The scale of each of these features is estimated by interpolation of the score of each feature.

After detecting feature points, a description vector is composed for each one. The description vector is a 512-bit binary code, which is based on comparing the brightness of the feature point with its neighbours. Then, in the end, the description vector in the reference image is compared with all of the description vectors of the sensed image through logical XOR operator. If the minimum error was lower than a predefined threshold *T*, then both features are considered as matched pairs [10].

Alahi et al. [12] introduced the Fast Retina Keypoint (FREAK) descriptor. This descriptor is similar to the BRISK but with a different sampling pattern which is inspired by the human visual system, the retina. In this sampling pattern, going away from the centre of the pattern, the standard deviation of the Gaussian function becomes larger and the density of pattern points becomes smaller. They showed that the FREAK algorithm requires less memory at runtime and is faster than the BRISK algorithm.

3. THE PROPOSED METHOD

BRISK and FREAK descriptors compare the brightness of the region around keypoints by using their pattern and build a binary code as output for each feature point. In this section, inspired by these two algorithms, an algorithm which is superior in discriminatory power of both algorithms is proposed. We call the proposed algorithm as DPML-RISK.

In the DPML-RISK algorithm, the keypoints are more robust to the scale, rotation, perspective and brightness changes than the BRISK algorithm.

The DPML-RISK algorithm is also a modular descriptor which can be used with any other scale invariant keypoint detection algorithms, which can estimate the scale factor of the keypoints.

3. 1. Keypoints Detection The algorithm proposed in this study uses SIFT algorithm for the keypoint detection. This detector utilizes DOG function, which is a good approximation of the normalized Laplacian function for the keypoints detection. At first, a scale space is formed from the image, and then the Gaussian Blur filter is used at each layer of the scale space [13]. In order to make the DOG functions, each blurred layer of scale space is subtracted from its above layer.

After making the scale space of DOG, points which are local extremums of the scale space are defined as the keypoints. Extremum point is a pixel in a scale space which is maxima or minima among its 26 neighbours (8 neighbours in the same layer, 9 neighbours in the above layer and 9 neighbours in the below layer). The keypoints coordinates and their scale factors are stored in a matrix to be used in the keypoint descriptor.

3. 2. Keypoint Descriptor DPML-RISK generates a multi-level vector in output, which is the result of comparing the brightness of pixels around the keypoints obtained in the keypoint detection stage. This descriptor assigns the main direction to each keypoint. This direction is used to normalize the descriptor algorithm and to make it robust against image rotation.

3. 2. 1. Sampling Pattern and Rotation Estimation Since our proposed method is based on the BRISK algorithm, in this section the main idea of this algorithm is briefly explained. The BRISK algorithm uses sampling pattern shown in Figure 1 to estimate the direction of keypoints. The proposed algorithm like the BRISK algorithm uses the sampling pattern shown in Figure 1 to estimate the direction of keypoints. The direction of keypoint is an estimation of the gradient of keypoint and indicates the direction of maximum brightness variations.

To avoid aliasing and to reduce sensitivity to noise, Gaussian smoothing algorithm is applied at each point of the sampling pattern [10]. Blue points of the sampling pattern in Figure 1, denote sampling locations and red circles around the blue points correspond to the standard deviation of the Gaussian kernel used for smoothing. Also p_i and σ_i determine coordinates and the standard deviation of blue point i, respectively.

The standard deviation is selected large enough to let the smoothing function have adequate overlap with the adjacent points. This technique will cause all of the pixels around the keypoints to contribute to the estimation of direction.

To estimate the direction of a keypoint, the centre of the sampling pattern is moved to the coordinates of the keypoint and its scale is changed by the scale factor of the keypoint. Then the gradient between two points p_i and p_j is estimated by the following equation:

$$g(p_j, p_i) = (p_j - p_i)\frac{(I(p_j,\sigma_j) - I(p_i,\sigma_i))}{\| p_j - p_i \|^2} \tag{1}$$

where $I(p_i,\sigma_i)$ is the smoothed intensity values at the point p_i with the standard deviation σ_i of the Gaussian smoothing filter. This equation denotes the direction and brightness changes between the two points p_i and p_j. Then a set A, which contains all sampling-point pairs, is defined as follows:

$$A = \left\{ (p_i, p_j) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i < N \wedge j < i \wedge i, j \in N \right\} \tag{2}$$

The total number of sampling-point pairs is calculated as:

$$|A| = \frac{N(N-1)}{2} \tag{3}$$

where N is the number of points of the sampling pattern. Now, a subset of the "long-distance" of the set A is defined as follows:

$$L = \left\{ (p_i, p_j) \in A \mid \| p_i - p_j \| > \delta_{\min} \right\} \subseteq A \tag{4}$$

This subset contains those sampling-point pairs in A, which their distance is more than the threshold $\delta_{min} = 13.67T$ where T is the scale factor of the keypoint. The number of sampling-point pairs in the subset L for the selected threshold $\delta_{min} = 13.67T$ is equal to 870. Now the gradient of all sampling-point pairs of L is calculated by Equation (1). Then, the average gradient, which is considered as the gradient of the keypoint, is calculated over all sampling-point pairs of L using the following equation:

$$g = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \frac{1}{k} \sum_{(p_i - p_j) \in L} g(p_i - p_j) \tag{5}$$

where k is equal to the length of L. The average gradient shows the direction of maximum brightness of the keypoint. Finally, the angle of the gradient vector in radians is computed as follows:

$$\alpha = \arctan(g_y, g_x) \tag{6}$$

In Figure 2, each keypoint that is found in the image is shown with an arrow. The size and direction of the arrows correspond to the scale factors and the direction of keypoints, respectively.

3. 2. 2. Building the Descriptor After obtaining the scale factor and the direction of keypoint, a new directional sampling pattern shown in Figure 3 is proposed to build the descriptor.

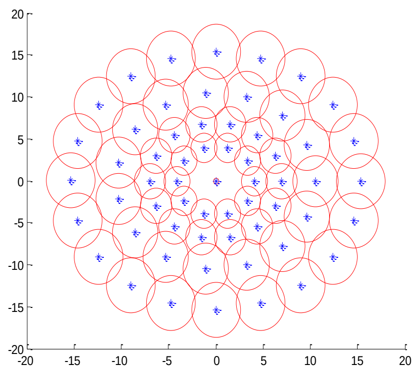
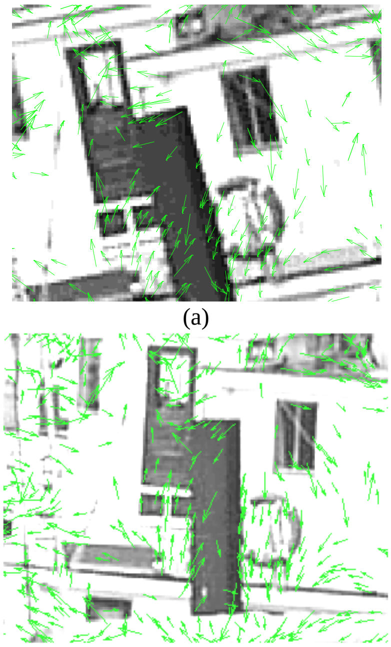


Figure 1. The sampling pattern which is used in the proposed algorithm to estimate the direction of keypoints



(a)

(b)

Figure 2. The location and direction of the keypoints which are found. The size and the direction of arrows correspond to the scale factors and the direction of keypoints respectively. (a) Sensed image. (b) Reference image

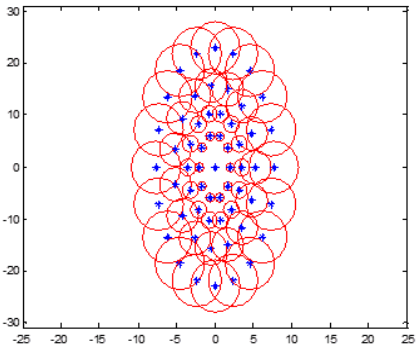


Figure 3. The directional sampling pattern used in the proposed algorithm to build the descriptor

The major axe of the directional pattern is approximately 2.5 times of the minor axe. To build the descriptor, we move the centre of the directional sampling pattern to the coordinates of keypoint, meanwhile scaling and rotating it by the scale factor and the direction of keypoint. Then a short-distance subset of set A is defined for the directional sampling pattern by the following equation:

$$S^d = \left\{ (p_i^d, p_j^d) \in A \mid \left\| p_i^d, p_j^d \right\| < \delta_{\max} \right\} \subseteq A \tag{7}$$

Superscript d denotes that the directional sampling pattern has been rotated. This subset contains all sampling point pairs in the directional sampling pattern that the distance between them is less than the threshold $\delta_{max} = 9.16T$. By changing the threshold level, δ_{max}, the length of the output code can be tuned. Then the brightness of each sampling-point pairs of set S^d is subtracted, and the function stated as follows:

$$\begin{aligned} & b = f\left([I(p_j^d, \sigma_j) - I(I(p_i^d, \sigma_i)], L\right) \\ & \forall (p_i^d, p_j^d) \in S^d \end{aligned} \tag{8}$$

which is used to map the result of this subtraction to a quantized number between -1 and 1. Here L is a user-defined even number representing the number of discretized levels. This function can be defined by the greatest integer number and the piecewise functions. If the number of levels increases, memory consumption of the generated code will increase. Figure 4 shows this function for $L = 6$.

This code describes a feature point based on the brightness of pixels which are inside of the pattern area. Figure 5 depicts the region around a keypoint in the proposed directional pattern and the simple pattern used in BRISK.

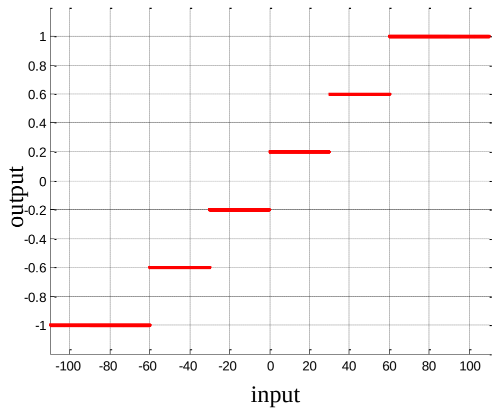


Figure 4. The encoding function which is used to encode brightness around keypoints with six levels

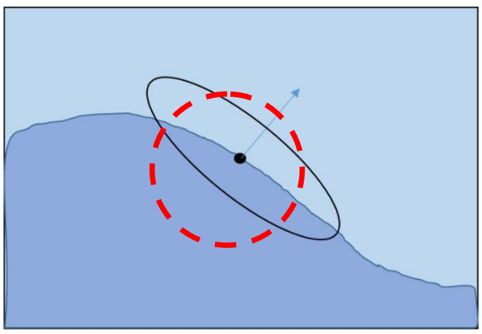


Figure 5. The regions around a keypoint that contribute building the description vector; (red-dashed line) the BRISK pattern and (black-solid line) the directional pattern

As it can be seen in Figure 5, more pixels of around edges than the pixels of background or foreground have been used to build the descriptor. Therefore, the proposed directional pattern provides more information about the edges that contribute in the descriptor vector.

Figure 6 shows the block diagram of the proposed algorithm. As it was mentioned the proposed algorithm contains two steps. In the first step, the feature points are obtained by DOG function. In the second step, the direction of the keypoints is estimated, and the descriptor vector is built using the proposed directional sampling pattern.

4. EXPERIMENTAL RESULTS

The new proposed algorithm is evaluated on the dataset in reference [14]. The dataset contains 6 series of images. Each series includes original image and five transformations, i.e., rotation, scale change, viewpoint change, image blur, JPEG compression and illumination applied to the original image separately. The original image of each series is shown in Figure 7 and Table 1 represents the details of the dataset briefly.

In this context two criteria, Recall and Precision, are used for evaluation.

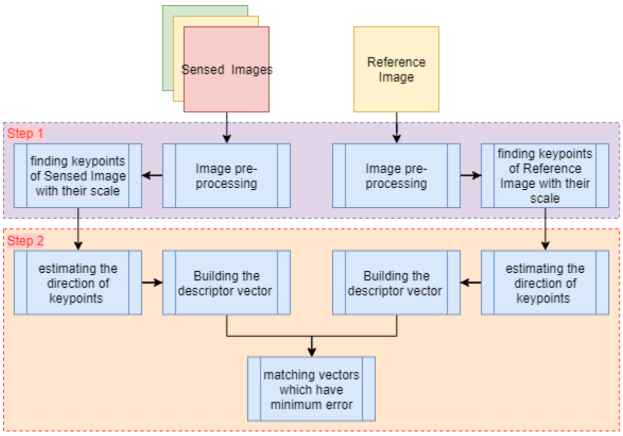


Figure 6. Block diagram of the proposed algorithm

Figure 7. Some original images of the dataset used for evaluation

TABLE 1. The details of the dataset

Series Name	Number of Images	Type of Transformations
bikes	6	image blurring
ubc	6	JPEG compression
boat	6	Rotation and scale
geaf	6	viewpoint changing
trees	6	blurring and rotation
leuven	6	Illumination changing

The Recall criterion is defined as the number of correct matching relative to the number of corresponding points in two images:

$$\text{Recall} = \frac{\#CorrectMatches}{\#Correspondences} \tag{9}$$

Corresponding points in two images depend on the overlap of images and repeatability of the detector algorithm. Since the recall measure shows the correct matching relative to the corresponding points of the algorithm, greater the amount of recall indicates the performance of the algorithm. The precision is another criterion that is used in the literature. This criterion is defined as the number of correct matched keypoints relative to the total number of matched keypoints.

$$\text{Precision} = \frac{\#CorrectMatches}{(\#CorrectMatches + \#FalseMatches)} \tag{10}$$

The precision is a number between zero and one. If the precision of an algorithm is one, it means that all points are properly matched; otherwise, if it is zero it means that all points are falsely matched. Sometimes the criterion of 1-precision is used instead of the precision. This criterion displays inaccuracy of the algorithm:

$$1\text{-Precision} = \frac{\#FalseMatches}{(\#CorrectMatches + \#FalseMatches)} \tag{11}$$

Figures 8 and 9 show the results of 30 tests on BRISK, FREAK, and DPML-RISK algorithms based on these two criteria. These figures show that the proposed DPML-RISK algorithm outperforms the BRISK algorithm. The DPML-RISK algorithm achieves the correct matched points almost two times higher than the BRISK algorithm and three times more than the FREAK. However, this improvement costs increase in memory consumption, while the DPML-RISK algorithm with $L=2$ in terms of memory and processing time similar to the BRISK algorithm.

It is worth mentioning that by increasing the index of images of the dataset, transformations are more destructive. Therefore, the number of correct matched points is decreased when the index of the image is increased.

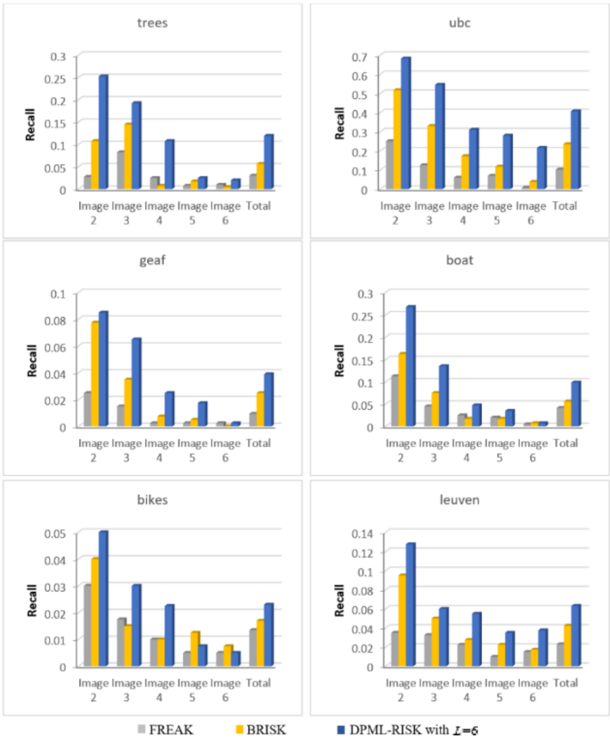


Figure 8. The results of the DPML-RISK, the FREAK, and the BRISK on the dataset

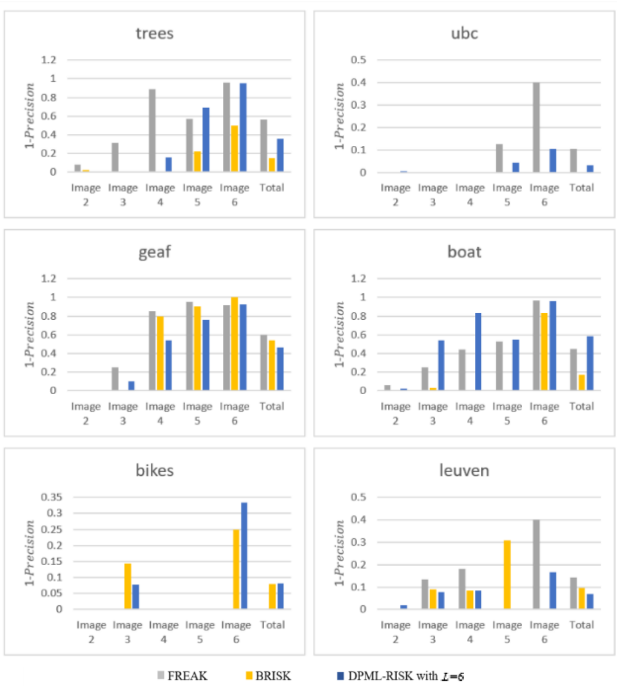


Figure 9. The results of the proposed algorithm, the FREAK, and the BRISK on the dataset

As Figure 8 demonstrates, the proposed algorithm finds twice corresponding points as much as the BRISK and the FREAK for images which are affected by rotation, scale, blurring and JPEG compression. In Bikes series the output is acceptable for lower light intensity. However by increasing the destruction, the corresponding points found in the proposed algorithm are comparable with the BRISK algorithm.

Also, Figure 9 shows the quantitative results for the 1-Precision criterion. The criterion for the proposed algorithm, shows an increase in comparison with the BRISK and virtually equal to the FREAK for all pictures of dataset.

Table 2 summarizes the results of the FREAK, the BRISK and the proposed algorithm on the given dataset which we performed with Features 2D class from OpenCV version 2.4.9 on the same hardware (Microsoft Surface Pro 3 / Intel Core i5 1.9 GHz Processor / 4 GB RAM).

Each row shows the sum of the true and false matching between image 1 and the transformed corresponding images in the dataset. Figure 10 shows the output of our algorithm in a pair of images on the dataset.

As it is expected, for all algorithms the number of true matches decreases respect to the increase in the destructiveness of the transformations. Also, the number of true matches of DPML-RISK algorithm in comparison with the FREAK and the BRISK algorithms has been increased by almost four and two times, respectively.

The number of vector levels, *L*, can be selected according to the availabe computing power. If *L* is high, the recall criterion of the proposed algorithm increases, despite the fact that proposed algorithm requires more computing power.

5. CONCLUSION

In this paper, a new algorithm called DPML-RISK was proposed. Fundamentally, the DPML-RISK algorithm is based on the keypoint detector of the SIFT algorithm.

TABLE 2. The results of the proposed algorithm, the FREAK and the BRISK algorithms

	FREAK		BRISK		DPML-RISK with *L*=6	
	#TP[a]	#FP[b]	#TP	#FP	#TP	#FP
All Images 2	203	4	417	1	611	5
All Images 3	131	25	263	3	428	70
All Images 4	65	112	104	13	239	118
All Images 5	47	39	81	48	167	70
All Images 6	20	214	33	58	117	297
Total	466	394	898	123	1562	560

[a] the number of true positive matches
[b] the number of false positive matches

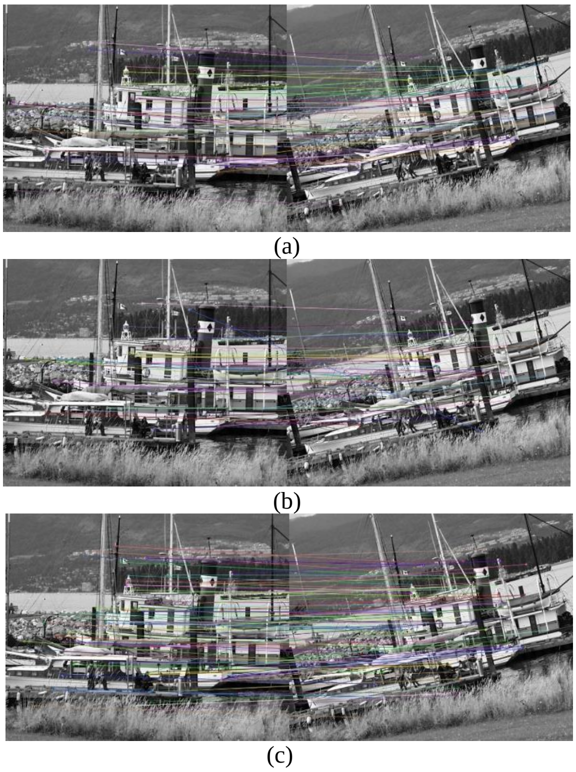


Figure 10. The result of three algorithms on the image 1 and 2 in boat series. (a) The result of BRISK. (b) The result of FREAK. (c) The result of proposed algorithm. As it can be seen, our algorithm could find coresponding point more than two other algorithms

Also, the sampling pattern for rotation estimation of the keypoints is based on the BRISK algorithm. It is recommended that use both the directional pattern and the multi-level codes in order to have a better description of the keypoints. The directional pattern led to derive more valuable information for the construction of descriptor. As a result, the algorithm has more ability to produce discriminative descriptor code for keypoints. The advantage of using a multilevel vector over the binary vector is not only for the more accurate magnitude of the point pairs, but also a larger number of useful points will also be involved in building codes. The significant results of the proposed algorithm are indebted to directional pattern and using multilevel codes.

6. REFERENCES

1. Oliveira, F.P., Tavares, J.M.R.J.C.m.i.b. and engineering, b., "Medical image registration: A review", Vol. 17, No. 2, (2014), 73-93.
2. Patel, Paresh M., and Vishal M. Shah. "Image registration techniques: a comprehensive survey." International Journal of Innovative Research and Development, Vol. 3, No. 3, (2014), 68-78.
3. Zitova, B., Flusser, J.J.I. and computing, v., "Image registration methods: A survey", Vol. 21, No. 11, (2003), 977-1000.
4. Karimi, M., A. Sadeghi Niaraki, and A. Hosseininaveh Ahmadabadian. "Automatic Recognition of Coded Targets Using Feature Based Matching Algorithms in a Ubiquitous GIS." Journal of Geomatics Science and Technology 7, No. 1 (2017): 1-13.
5. Hassaballah, M., Aly Amin Abdelmgeid, and Hammam A. Alshazly. "Image features detection, description and matching." In Image Feature Detectors and Descriptors, pp. 11-45. Springer, Cham, 2016.
6. Lowe, D.G.J.I.j.o.c.v., "Distinctive image features from scale-invariant keypoints", Vol. 60, No. 2, (2004), 91-110.
7. Lowe, David G. "Object recognition from local scale-invariant features." In Computer vision, 1999. The proceedings of the seventh IEEE international conference on, vol. 2, pp. 1150-1157. IEEE, 1999.
8. Bay, Herbert, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. "Speeded-up robust features (SURF)." Computer vision and image understanding, Vol. 110, No. 3 (2008): 346-359.
9. Khatami, A., Babaie, M., Tizhoosh, H., Nazari, A., Khosravi, A. and Nahavandi, S.J.I.J.o.E.-T.C.A., "A radon-based convolutional neural network for medical image retrieval", Vol. 31, No. 6, (2018), 910-915.
10. Leutenegger, S., Chli, M. and Siegwart, R.Y., "Brisk: Binary robust invariant scalable keypoints", in Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE. Vol., No. Issue, (2011), 2548-2555.
11. Mair, E., Hager, G.D., Burschka, D., Suppa, M. and Hirzinger, G., "Adaptive and generic corner detection based on the accelerated segment test", in European conference on Computer vision, Springer. Vol., No. Issue, (2010), 183-196.
12. Alahi, A., Ortiz, R. and Vandergheynst, P., "Freak: Fast retina keypoint", in 2012 IEEE Conference on Computer Vision and Pattern Recognition, Ieee. Vol., No. Issue, (2012), 510-517.
13. Lindeberg, T.J.E.o.M., "Scale-space theory", Vol., No., (2001).
14. Mikolajczyk, K., Schmid, C.J.I.t.o.p.a. and intelligence, m., "A performance evaluation of local descriptors", Vol. 27, No. 10, (2005), 1615-1630.

DPML-Risk: An Efficient Algorithm for Image Registration

S. Kazemi, M. R. Ahmadzadeh

Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

PAPER INFO

Paper history:
Received 16 March 2018
Received in revised form 30 August 2018
Accepted 26 October 2018

Keywords:
Feature Detection
Binary Descriptor
Image Registration
Scale Invariant Feature Transform

چکیده

تطبیق و ره‌گیری اهداف در تصاویر نقش عمده‌ای را در صنعت ایفا می‌کنند. یکی از روش‌های تطبیق استفاده از الگوریتم‌های مبتنی بر ویژگی است که به‌طورمعمول در دو مرحله صورت می‌گیرد. در اولین مرحله‌ی این الگوریتم‌ها، ویژگی‌هایی در تصویر دریافتی و تصویر هدف با استفاده از فضای مقیاسی استخراج می‌شود. سپس برای هر نقطه کلیدی به‌دست‌آمده در مرحله اول بر اساس شدت روشنایی اطراف آن یک بردار توصیفگر ساخته می‌شود. در این مقاله الگوریتم جدیدی با الهام از دو الگوریتم BRISK و SIFT، پیشنهادشده است. این الگوریتم از الگوی جهتی برای توصیف ویژگی استفاده می‌کند. جهت این الگو در جهت عمود بر زاویه ویژگی قرار می‌گیرد که این امر باعث می‌شود که اطلاعات مفیدتری از روشنایی‌های اطراف نقاط کلیدی در ساخت بردار توصیفگر استفاده شود. علاوه بر این بردار خروجی در الگوریتم پیشنهادی به‌صورت یک بردار چند سطحی است که باعث قدرت تمایز بیشتری می‌شود. نتایج این پژوهش نشان می‌دهد که قدرت تمایز و پایداری الگوریتم پیشنهادی نسبت به الگوریتم BRISK قوی‌تر شده است و از نظر کارآمدی الگوریتم تا حدود زیادی با الگوریتم BRISK برابر است.

***doi**: 10.5829/ije.2018.31.11b.09*