

Research Article

Random Binary Local Patch Clustering Transforms Based Image Matching for Nonlinear Intensity Changes

Han Wang ¹, Zhihuo Xu,¹ and Hanseok Ko ²

¹*School of Transportation, Nantong University, 9 Seyuan Rd, Nantong, China*

²*School of Electrical Engineering, Korea University, Seoul, Republic of Korea*

Correspondence should be addressed to Han Wang; hanwang@ntu.edu.cn

Received 25 April 2018; Revised 10 August 2018; Accepted 26 August 2018; Published 19 September 2018

Academic Editor: Konstantinos Karamanos

Copyright © 2018 Han Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a new feature descriptor that is suitable for image matching under nonlinear intensity changes. The proposed approach consists of the following three steps. First, a binary local patch clustering transform response is employed as the transform space. The value of the new space exhibits a high similarity after changes in intensity. Then, a random binary pattern coding method extracts raw feature histograms from the new space. Finally, the discrimination of the proposed feature descriptor is enhanced by using a multiple spatial support region-based binning method. Experimental results show that the proposed method is able to provide a more robust image matching performance under nonlinear intensity changes.

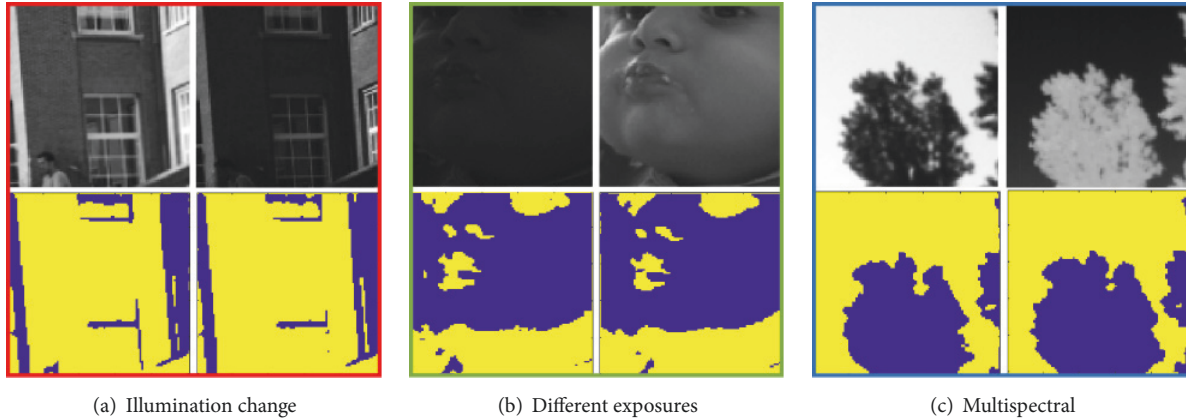
1. Introduction

With the rapid development of modern sensors and big data technology, image matching has been applied to many computer vision applications, such as multiple sensor fusion [1, 2], 3D reconstruction [3], and depth information estimation [4, 5]. Nevertheless, when the intensity value of an image pair contains an obvious nonlinear variation, image matching becomes very challenging. Typical intensity changes include changes in illumination [6, 7], different foci [8], flash versus no-flash [9], multispectral images [10, 11], multimodalities [12], and blurred images [13]. The most direct approach to this challenging task is to construct a feature descriptor that is effective for different types of image transformations.

Image matching under illumination changes has a very important effect on the performances of visual object detection, tracking, and recognition. The local relationship between the intensity values of pixels is employed to construct an illumination-independent feature descriptor, such as a local binary pattern (LBP) [14] and relative family features, namely, center symmetric local binary pattern (CS-LBP) [15] and center symmetric local trinary pattern (CS-LTP) [16]. Local features extracted from intensity rankings rather than raw intensity have become a more popular method recently

because intensity order distributions are invariant to monotonic intensity changes [17]. In this regard, Tang proposed an ordinal spatial intensity distribution (OSID) [17], which uses the relative order of pixel intensity within a patch to generate illumination-robust features. However, ordering pixels by discrete intensity yields significantly different order distributions when the illumination changes, thereby leading to changes in the nonlinear intensity. To address this problem, the exact order based descriptor (EOD) employed by Kim [18] used an exact order method.

Image matching in multispectral cases is more challenging than in cases of illumination change. Long wavelength infrared (LWIR) intensity variations are related to variations in the temperatures of objects, while variations in the RGB intensities come from colored objects and light reflections. Due to their different natures, the nonlinear relationship of the intensity values across these image types significantly complicates image matching due to the resulting lack of correlation between their respective gradients. Furthermore, LWIR images appear smoother, with losses of detail and texture [19], as Figure 1(c) shows. Thus, in multispectral cases, many novel feature descriptors, such as the scale invariant feature transform SIFT [20], lose their effectiveness. Inspired by SIFT, many SIFT-like descriptors have been



(a) Illumination change

(b) Different exposures

(c) Multispectral

FIGURE 1: Example of the binary local patch clustering transform results with (a) illumination change, (b) different exposures, and (c) multispectral images.

proposed for RGB/LWIR image matching. Cristhian [19] describes the Edge Oriented Histogram (EOH) descriptor, which is robust for characterizing multispectral key-points. Additionally, Mouats employs multioriented and multiscale Log-Gabor filters instead of multioriented Sobel descriptors in order to construct a new and upgraded feature descriptor PCEOH [21].

Local self-similarity (LSS) [22] is another novel method for feature extraction that captures the local internal layouts of self-similarities in the image. Inspired by LSS, Heinrich proposed a modality independent neighbourhood descriptor (MIND) [23] for multimodal image matching. Kim proposed a dense adaptive self-correlation (DASC) [24] with a series of adaptive self-similarity measures between patches sampled by randomized receptive field pooling, for which the sampling pattern was obtained through discriminative learning. However, every image patch has separate internal layout information. Despite employing a machine-learning algorithm [25], the fixed sampling pattern based self-similarity group feature (DASC) could not achieve perfect performance across all test images.

This paper presents a new descriptor called a random binary pattern of patch clustering (RBPPC), which employs binary local patch clustering transform responses as the transform space of the input image. Then, groups of random sampling responses in the new space are converted into different patterns as the “cluster bin” of the proposed histogram feature descriptor. Finally, multiple spatial support regions are designed in order to enhance the discrimination of the proposed feature descriptor.

This paper is organized as follows. Section 2 introduces the proposed RBPPC feature descriptor. Section 3 describes the experimental process and results analysis. Finally, the concluding remarks are given in Section 4.

2. Proposed Feature Descriptor

2.1. Binary Local Patch Clustering Transform. Instead of constructing a descriptor with a series of self-similarity

measures directly, in this paper a binary local patch clustering transform is proposed as the transform space, which performs robustly to intensity changes. The typical intensity changes include changes in illumination, different exposures, and multispectral images. For the multispectral images (e.g., RGB/LWIR images) in particular, there is not only little texture information but also low contrast in the LWIR image compared to the RGB image. Therefore, a binary local path clustering transform can generate a more similar transform space.

Figure 1 shows an example of the response of the binary local patch clustering transform to intensity changes. The first row of Figure 1 shows the input image pairs, featuring different types of intensity changes, and the second row shows the response of the image pairs to the binary local patch clustering transform. This example confirms that the new transformed space can be less sensitive to intensity changes than the raw intensity space. Therefore, building a robust intensity change feature descriptor requires an effective matching approach that confines the feature values of the transformed space in response to the changes in intensity.

The proposed binary local patch clustering transform consists of the following five steps:

(1) Extract the feature vector sets with identical orders from the source and reference images by employing a sliding window in the same direction (from left to right and up to down) as shown in Figure 2:

$$\begin{aligned} \mathbf{F}_S &= \{P_1, P_2 \cdots, P_N\} \\ \mathbf{F}_R &= \{Q_1, Q_2 \cdots, Q_N\}, \end{aligned} \quad (1)$$

where F_S and F_R are the local feature vector sets of the source and reference images, respectively. P_i and Q_i are the feature vectors, which consist of the intensity values of the i -th corresponding local image patches of the source and reference images. N is the total number of local feature vectors, as Figure 2 Step (2) shows.

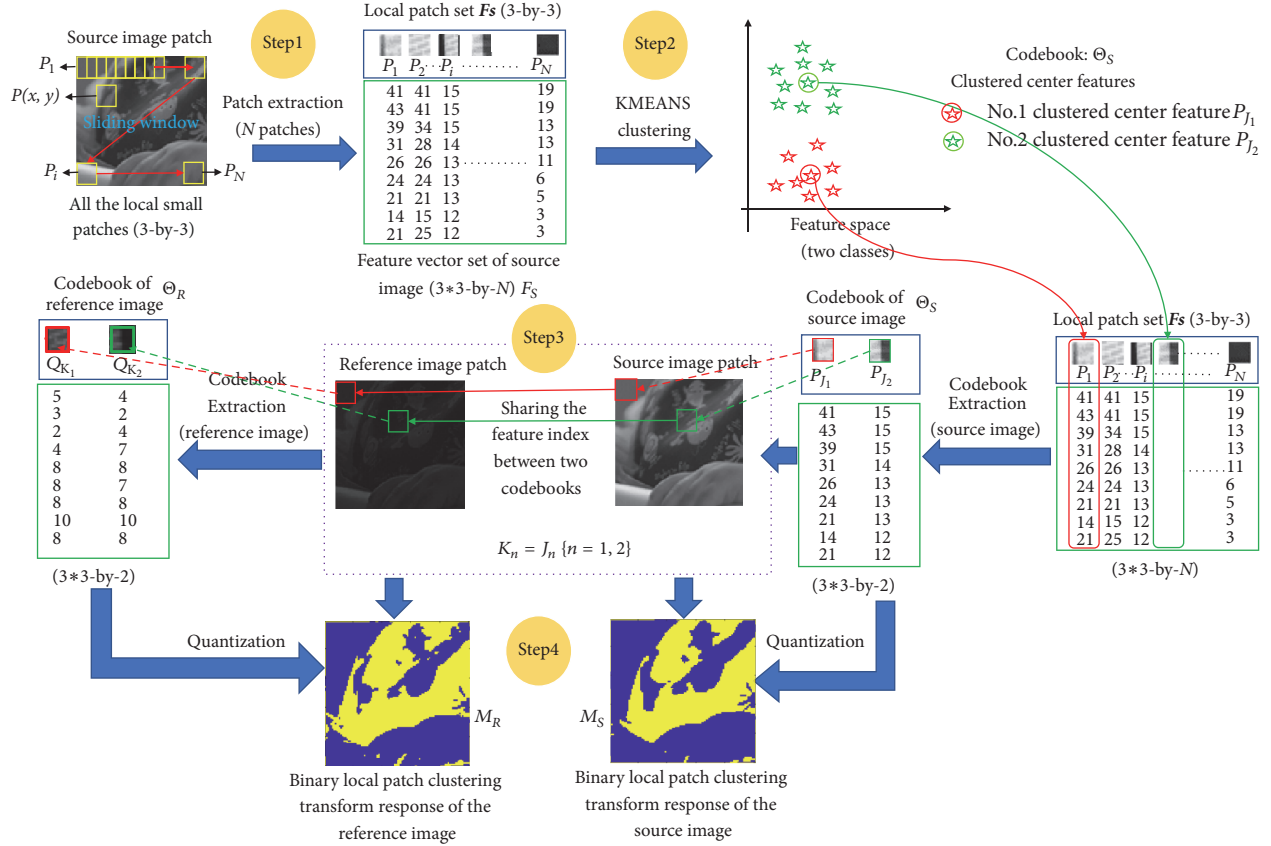


FIGURE 2: Binary local patch clustering distribution building process.

(2) Build a codebook of the source image using a K-means clustering algorithm [26] from F_S :

$$\theta_S = \{P_{j_1}, P_{j_2}\}, \quad (2)$$

where θ_S represents the codebook of the source image, which consists of two feature vectors. P_{j_1} and P_{j_2} represent the binary clustering centers extracted by K-means from F_S . Therefore, P_{j_1} and P_{j_2} are actually the j_1 -th and j_2 -th feature vectors in F_S .

(3) Build a codebook of the reference image by sharing the index information of the source image binary clustering centers P_{j_1} and P_{j_2} :

$$\theta_R = \{Q_{k_1}, Q_{k_2}\} \quad (3)$$

$$k_n = j_n \quad \{n = 1, 2\},$$

where θ_R represents the codebook of the reference image and j_n is the codebook feature vector index of the corresponding local patches in the source local feature set F_S . The relationship established in (3) depicts the source and reference codebooks that share the same clustering index ($k_n = j_n$). Therefore, although the codebooks P_{k_n} and Q_{j_n} consist of different intensities, they represent the n -th local patch clustering centers of the source image and reference image, respectively. This simple method deals effectively with the quantization problem caused by intensity differences.

(4) Generate the binary local patch clustering transform response using codebook-based quantization:

$$M_S(x, y) = \arg \min_j \{SSD(P(x, y), \theta_S(j))\} \quad (4)$$

$$M_R(x, y) = \arg \min_k \{SSD(Q(x, y), \theta_R(k))\},$$

where M_S and M_R are the binary local patch clustering transform responses of the source and reference images, respectively. SSD presents the sum of the square differences of the intensity values $P(x, y)$ and $Q(x, y)$ representing the intensity features of the local patch located at the (x, y) points in the source image patch and the reference image, respectively. As a result, either the source or the reference local patches are assigned a binary class label. Figure 2 shows the entire processing of the binary local patch clustering transform algorithm.

In order to illustrate the advantages of the proposed binary local patch clustering transform algorithm (BLPCT), three classical image binarization algorithms, namely, the Otsu method, the Kittler method, and the Niblack method, are employed as comparison objects. The comparison experiment is set up as follows: 100 local image patch pairs under intensity changes with different resolutions are randomly selected from the DB1 illumination changes database (Leuven), as the first row of Figure 3 shows. Then, a similarity test

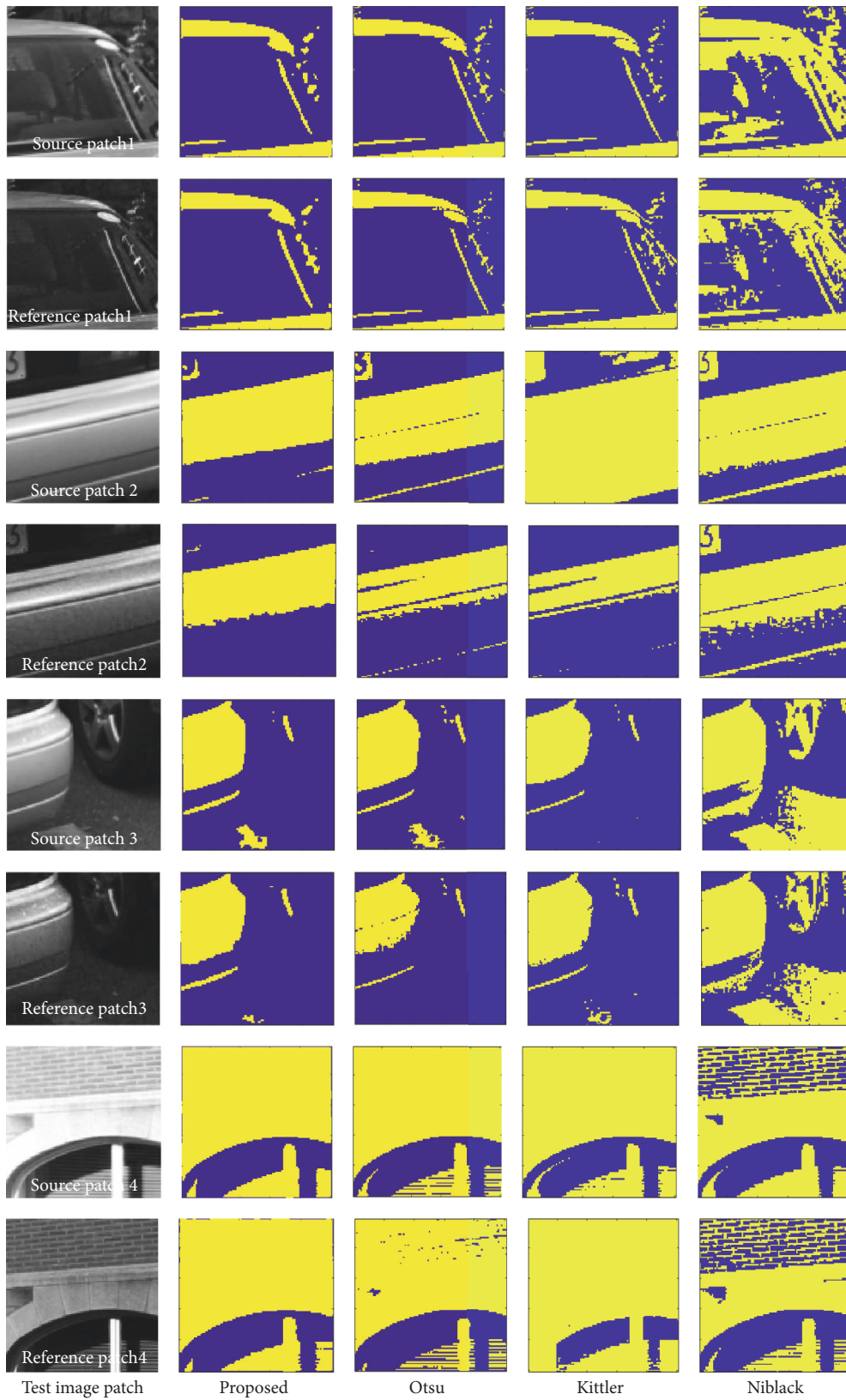


FIGURE 3: Comparison performance of the different image binarization methods.

TABLE 1: Comparison experimental results of different binarization methods.

	Proposed	Otsu	Kittler	Niblack
Average Similarity	98.33%	90.31%	88.42%	90.14%
Average Running Time	0.029 Sec	0.026 Sec	0.012 Sec	0.65 Sec

TABLE 2: Examples of the comparison experimental results of different image binarization methods.

	Proposed	Otsu	Kittler	Niblack
Test patch pair 1	98.33%	96.83%	88.42%	97.32%
Test patch pair 2	92.54%	91.77%	43.48%	97.09%
Test patch pair 3	97.85%	93.59%	96.26%	91.87%
Test patch pair 4	99.60%	94.68%	90.54%	91.91%
Average similarity	97.31%	91.72%	91.72%	89.54%

of patch pair binarization results, S_{im} , is employed as an evaluation function (5) to evaluate the effectiveness of different binarization algorithms. Finally, the average similarity of 100 image patch pair binarization results under intensity is used as a qualitative evaluation indicator of their performances.

$$Sim = \frac{\text{number of identical pixels}}{\text{total number of pixels}} \times 100\%. \quad (5)$$

Table 1 shows the comparison experimental results of the proposed method and the Otsu, Kittler, and Niblack methods. The first row of Table 1 contains the average similarity of different image binarization results for 100 random image patch pairs under intensity changes. The second row of Table 1 contains the average running time of the different image binarization algorithms. From Table 1, we can clearly see that the average similarity of the image binarization results for the proposed method is obviously higher than the average similarity for other classical image binarization algorithms, which illustrates that the proposed BLPCT algorithm performs more effectively under intensity changes compared with other conventional methods. Moreover, the computational complexity is similar to the Otsu algorithm.

Figure 3 shows some experimental result examples of the proposed binary local patch clustering transform (BLPCT) and the classical Otsu, Kittler, and Niblack binarization algorithms. Specifically, the first column of Figure 3 shows the test image patch pair under intensity changes (the source image patch and the reference image patch). The second column of Figure 3 shows the proposed BLPCT responses of the test image patch pair, and the third column of Figure 3 shows the Otsu algorithm responses of the test image patch pair. The fourth column shows the Kittler binarization algorithm results of the test image patch pair, and the fifth column shows the Niblack binarization algorithm results of the test image patch pair.

The corresponding evaluation results are summarized in Table 2. It is clear that the proposed binary local patch clustering transform algorithm performs more effectively than other classical image binarization algorithms under intensity changes, since the average similarity of the BLPCT

responses under intensity changes is 6–7% higher than that of other binarization algorithm responses.

In theory, according to the definition of the classical Otsu algorithm [27], the image histogram is separated into two groups using a gray value. When the variance of the two divided groups becomes a maximum, the gray value is seen as the best threshold. The algorithm utilizes the category variance as the criterion and chooses the gray value as the optimal threshold, which is the maximum of the variance between the categories and the minimum of variance within the categories. The Kittler algorithm [28] is also a gray level threshold based image binarization method, whose goal is to find the minimum error threshold. The Niblack algorithm [29] employs the local mean and the standard deviation to estimate the optimal threshold.

The proposed BLPCT algorithm is based on the local image patch classification result. Compared with a global or local intensity threshold, a local image patch can provide more information, such as the neighbouring intensity relationship and the spatial position information, which is more important to intensity changes.

2.2. Random Binary Pattern Coding Strategy. In order to convert the response of a binary local patch clustering transform to a feature histogram, this paper proposes a random binary pattern coding method, defined as follows:

$$F(\Phi_j(\mathbf{I})) = \text{pattern index} \quad (6)$$

$$\Phi_j(\mathbf{I}) = \{\phi_j^1(\mathbf{I}), \phi_j^2(\mathbf{I}), \phi_j^3(\mathbf{I}), \phi_j^4(\mathbf{I})\},$$

where $F(\Phi_j(I))$ represents the binary pattern index of the j -th random sampling response of image I . $\Phi_j(I)$ is a random binary pattern that consists of four sampling point responses of image I ($\Phi_{j1}(I)$, $\Phi_{j2}(I)$, $\Phi_{j3}(I)$ and $\Phi_{j4}(I)$).

Mathematically, $\Phi_j(I)$ presents a permutation of four random sample point responses to input image I . Therefore, given that the random sampling result of a binary local patch clustering transform response is either 0 or 1, when applying $\Phi_j(I)$ (6) to the binary local path clustering transform

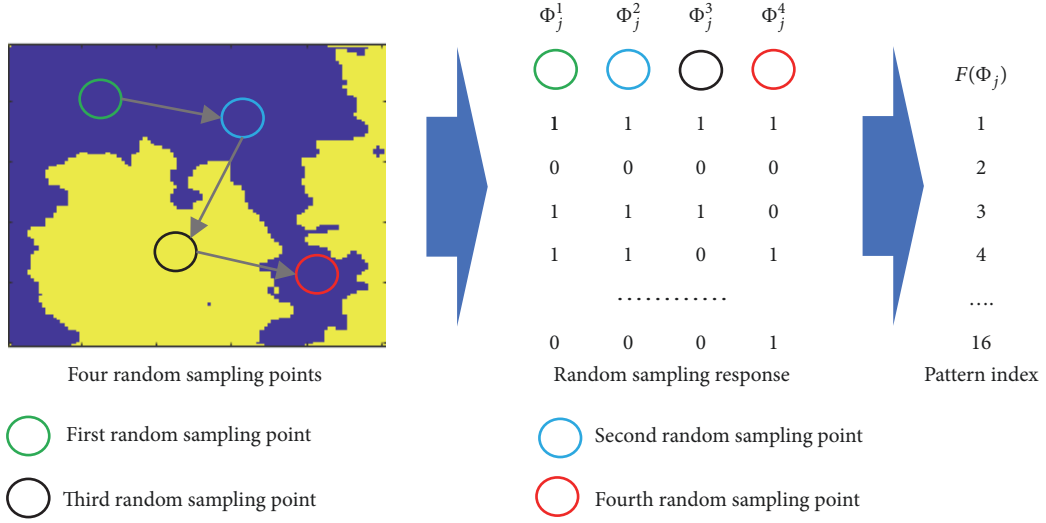


FIGURE 4: Illustration of the random binary pattern coding method.

TABLE 3: $F(\Phi_j(M))$ of the binary local patch clustering transform response.

$\Phi_j^1(M)$	$\Phi_j^2(M)$	$\Phi_j^3(M)$	$\Phi_j^4(M)$	$F(\Phi_j(M))$
1	1	1	1	1
0	0	0	0	2
1	1	1	0	3
.....
0	0	0	1	16

response M , the complete permutation of $\Phi_j(M)$ consists of 16 different patterns, as shown in Table 3.

The process of the random binary pattern coding method is explained clearly in Figure 4.

With the pattern index in Table 3, a feature mapping function π is defined to map the $\Phi_j(I)$ permutation into a 16-dimensional feature vector V_{16}^i , whose elements are all 0 except for the i th element, which is 1. The mathematical definition of π is

$$\pi(F(\Phi_j(\mathbf{I}))) = V_{16}^i = \left\{ 0, \dots, 0, \underset{i=F(\Phi_j(\mathbf{I}))}{1}, 0, \dots, 0 \right\}. \quad (7)$$

Therefore, based on the above definitions, the histogram feature of the proposed RBPPC can be defined by

$$\mathbf{H}_{\text{RBPPC}}(\mathbf{M}) = \sum_{j=1}^n \pi(F(\Phi_j(\mathbf{M}))), \quad (8)$$

where M represents the binary local patch clustering transform response and n represents the random sampling time. Figure 5 illustrates the histogram feature of the RBPPC generation process. First, four points are randomly sampled n times and the computing pattern index is calculated according to (6) for each sampling response. Then, the HRBPPC histogram is constructed by counting the pattern index number with reference to (8).

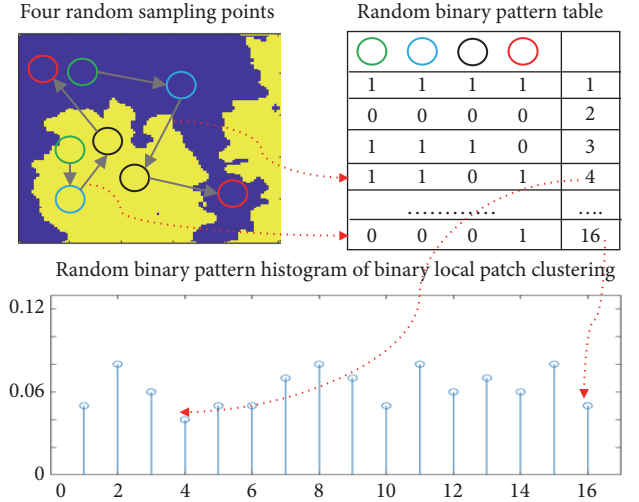


FIGURE 5: Illustration of the histogram feature of the RBPPC generation process.

2.3. Chamfer Distance-Based Spatial Constraint. With the help of the proposed binary local patch clustering transform, the transform space exhibits high similarity compared to the original image space under intensity changes, as shown by the above examples. However, there are still some tiny differences between the transform responses of the source and reference images. In other words, it is difficult for the binary local patch clustering transform to provide an absolutely identical transform space under intensity changes, such as those shown in Figures 7(c) and 7(d).

Therefore, in order to extract features that are more similar from the binary local patch clustering transform space under intensity changes, the Chamfer distance [30] response is applied in order to provide a spatial constraint for the RBPPC descriptor extraction. This spatial constraint can effectively enhance the robustness of RBPPC descriptor. The

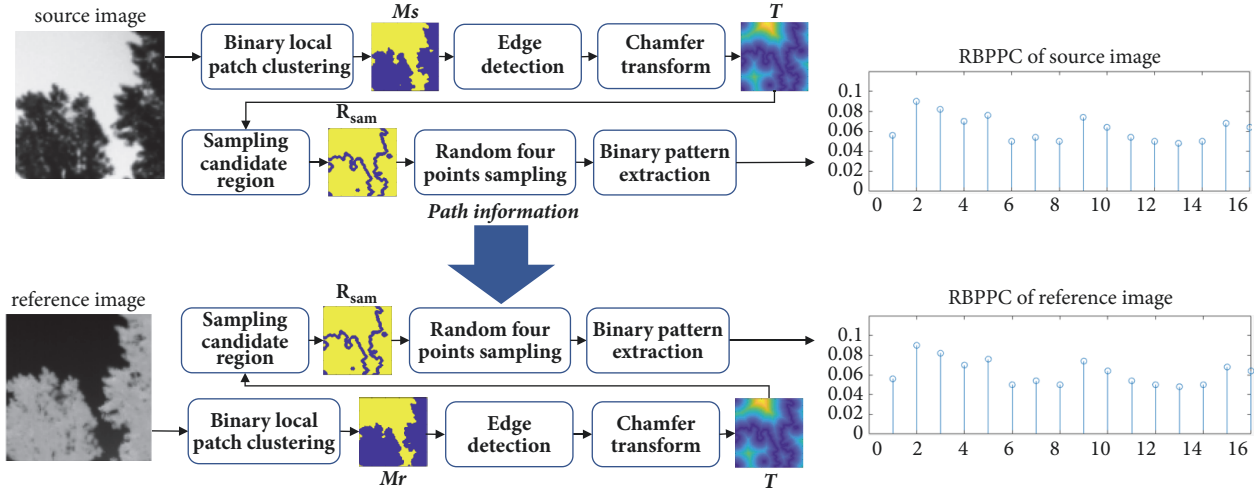
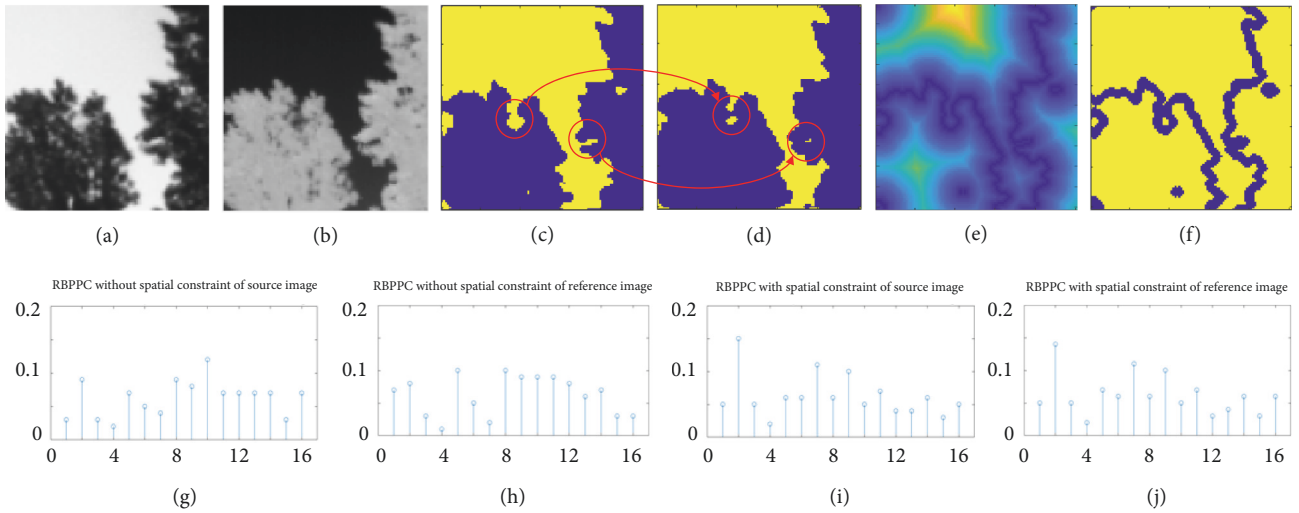


FIGURE 6: Example of the spatial constraint enhanced RBPPC histogram extraction process.


 FIGURE 7: Example of the RBPPC under different generation conditions for a multispectral image pair: (a) RGB source image, (b) LWIR reference image, (c) M_s from the source image (a), (d) M_r from the reference image (b), (e) Chamfer distance response of (a), (f) spatial sampling point candidate regions, (g) RBPPC of the source image, (h) RBPPC of the reference image, (i) RBPPC of the source image with the spatial sampling point candidate regions, (j) RBPPC of the reference image with the spatial sampling point candidate regions.

detailed process is described as follows: First, the Chamfer distance map is extracted from the binary local patch clustering transform of the source image as shown in Figure 7(e).

$$T = \text{Chamfer}(M_s), \quad (9)$$

where T is the Chamfer distance map and M_s presents the local patch clustering transform response of the source image.

Then, the spatial sampling point candidate regions are built from the Chamfer distance map T by a distance threshold D_{th} as follows:

$$R_{sam} = \{(x, y) \mid (x, y) \in T \ \& \ T(x, y) > D_{th}\}, \quad (10)$$

where R_{sam} represents the spatial sampling point candidate regions that are shown by the yellow region of Figure 7(f). Figure 6 shows the entire spatial constraint enhanced RBPPC

histogram feature extraction process. We can clearly see that, with the help of the sampling candidate region R_{sam} , the RBPPC histogram features under intensity changes are more similar.

Figure 7 shows an example of the RBPPC descriptor with an RGB/LWIR image pair under different generation conditions. Figures 7(g) and 7(h) show RBPPC histograms of input image pairs without spatial sampling point candidate region constraints, and there are obvious differences between those two histograms. In contrast, as Figures 7(i) and 7(j) show, when using spatial sampling point candidate region, the RBPPC feature pairs extracted from the source and reference images exhibit high similarity.

Moreover, Figure 8 shows an example of the RBPPC descriptor with a different exposure image pair under different generation conditions and Figure 9 shows an example

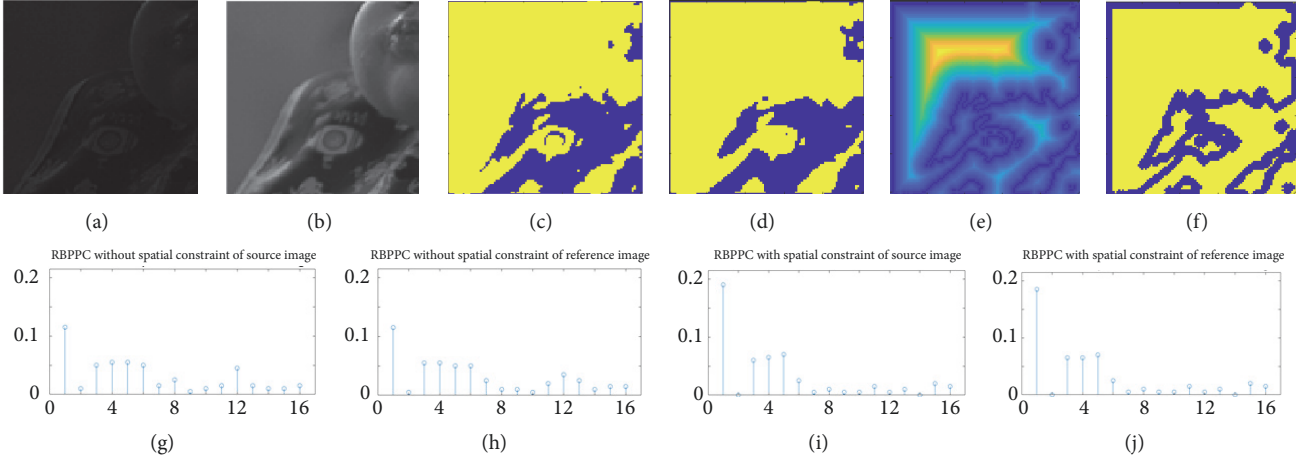


FIGURE 8: Example of the RBPPC under different generation conditions for an image pair with different exposures: (a) source image, (b) reference image, (c) M_s from the source image, (d) M_R from the reference image (b), (e) Chamfer distance response of (a), (f) spatial sampling point candidate regions, (g) RBPPC of the source image, (h) RBPPC of the reference image, (i) RBPPC of the source image with the spatial sampling point candidate regions, (j) RBPPC of the reference image with the spatial sampling point candidate regions.

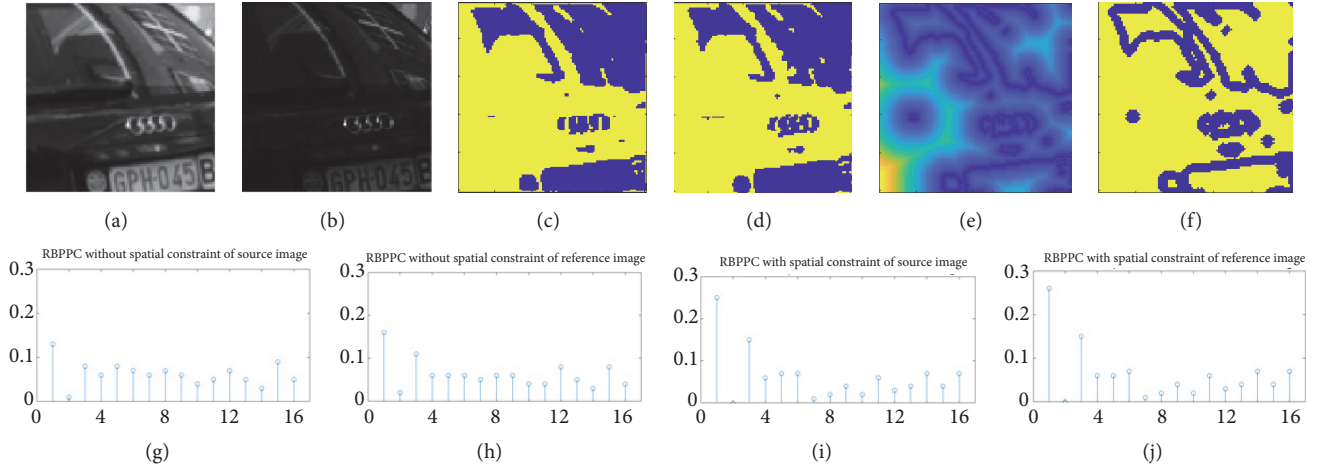


FIGURE 9: Example of the RBPPC under different generation conditions for an image pair with different illuminations: (a) source image, (b) reference image, (c) M_s from the source image, (d) M_R from the reference image (b), (e) Chamfer distance response of (a), (f) spatial sampling point candidate regions, (g) RBPPC of the source image, (h) RBPPC of the reference image, (i) RBPPC of the source image with the spatial sampling point candidate regions, (j) RBPPC of the reference image with the spatial sampling point candidate regions.

of the RBPPC descriptor with an image pair featuring different illuminations. The second rows of Figures 7, 8, and 9 feature different RBPPC descriptor generation conditions with and without spatial sampling point candidate region constraints. These examples demonstrate that the similarity of the proposed RBPPC descriptors for image pairs under intensity change is markedly improved by using the Chamfer distance transform based spatial sampling point candidate regions.

2.4. Multiple Spatial Support Regions Binning Method. In order to enhance the discrimination of the proposed RBPPC feature descriptor, this study applies a multiple spatial region-based binning method. Figure 10 shows 17 different spatial

support region binary masks designed to enhance the discrimination of the proposed RBPPC descriptor. The white parts of the spatial support region masks represent 1 while the black parts of spatial support region masks represent 0.

The multiple spatial support region-based RBPPC histogram feature descriptor vectors are defined as follows:

$$\mathbf{H}_{\text{RBPPC}}^{\text{Multi}} = \{\mathbf{H}_{\text{RBPPC}}^1, \mathbf{H}_{\text{RBPPC}}^2, \dots, \mathbf{H}_{\text{RBPPC}}^{17}\}$$

$$\mathbf{H}_{\text{RBPPC}}^i = \sum_{j=1}^n \pi_{\mathbf{X} \in \mathbf{C}_i} (F(\Phi(\mathbf{X}))) \quad (11)$$

$$\mathbf{C}_i = \mathbf{R}_{\text{Sup}}^i \otimes \mathbf{R}_{\text{Sam}},$$

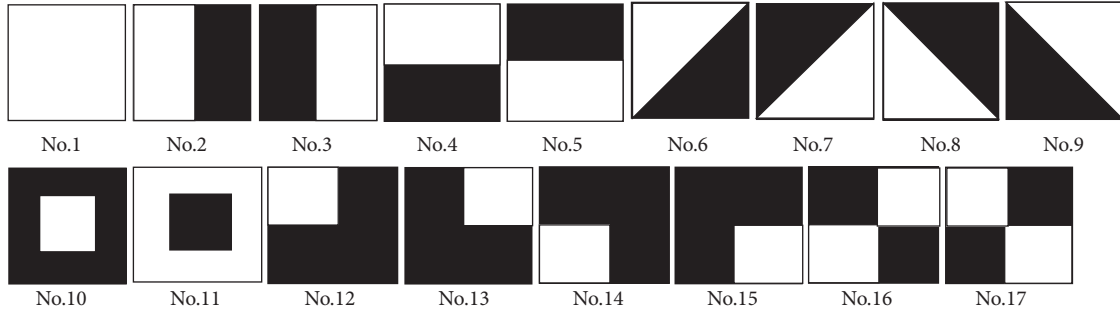


FIGURE 10: 17 different spatial support region binary masks.

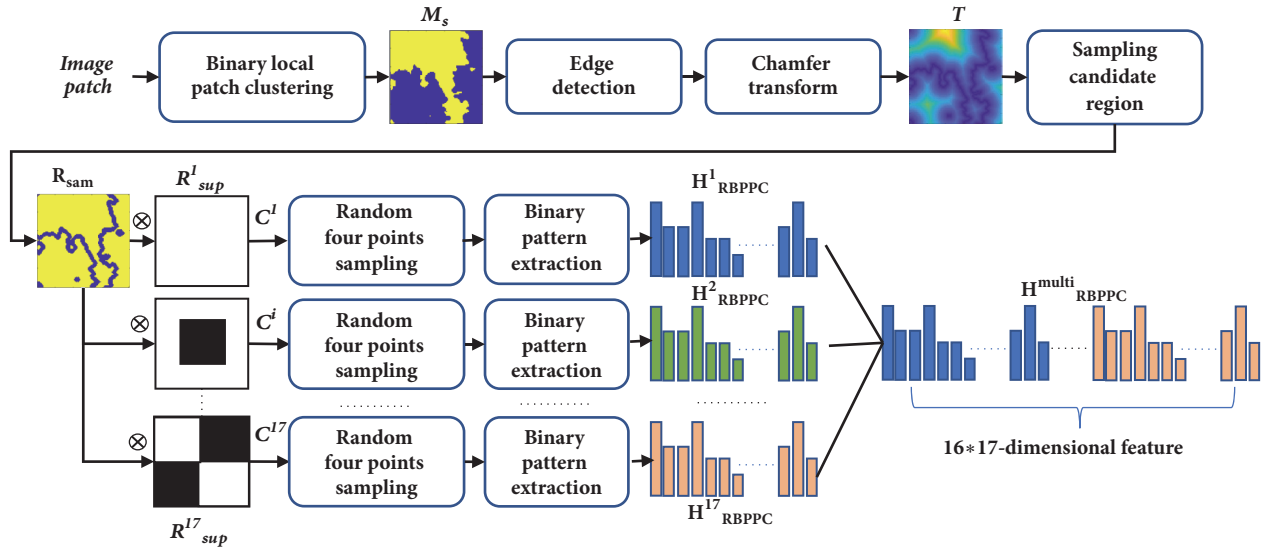


FIGURE 11: The flowchart of the feature extraction of multiple spatial support region-based RBPPC histograms.

where H^i_{RBPPC} represents the RBPPC histogram of i -th spatial support region. H^{Multi}_{RBPPC} represents the multiple spatial support region-based RBPPC histogram feature descriptor vectors, which consist of 17 H^i_{RBPPC} . π is a mapping function, which is defined to map a pattern index $F(\Phi_j(X))$ into a 16-dimensional feature vector V_{16}^i , as (7) shows.

Moreover, C_i represents the actual spatial random point sampling candidate regions of i -th actual spatial support region for the binary local patch clustering transform response M , and C_i is generated with the following two components, as (11) shows:

- (1) the binary mask of the i -th spatial support region R^i_{sup} , which is shown in Figure 10;
- (2) the spatial sampling point candidate regions R_{sam} , which are shown by the yellow regions of Figure 11.

Therefore, the multiple support region-based RBPPC has (16×17) 272-dimensional feature descriptor vectors. The flowchart of the feature extraction of multiple spatial support region-based RBPPC histograms is shown in Figure 11. Figure 12 shows examples of the multiple support region binning method enhanced RBPPC feature descriptor, used with different types of image pairs.

3. Experimental Results

We implemented the proposed algorithm with MATLAB 2016b software and a desktop computer with a 3.30 GHz Intel 5 processor. Three popular published image database sets under intensity changes were employed as test data, which consisted of illumination changes [31], different exposures [8], and multispectral images (RGB/LWIR) [32]. These datasets all provided the ground truth information about key-point positions and their corresponding matching results.

To be specific, the illumination changes database (Leuven) is a standard Oxford image matching dataset with natural illumination changes, which consists of six outdoor scene images under various degrees of illumination changes. These images are employed in this paper as the image matching test DB1 (illumination changes), which is shown in Figure 13(a). The second image matching test DB2 (different exposures) consists of four image pairs with obviously different exposures, and three image pair examples are shown in Figure 13(c). The third image matching test DB3 (multispectral images) consists of 44 multispectral image pairs (any image pair including an RGB image and an LWIR

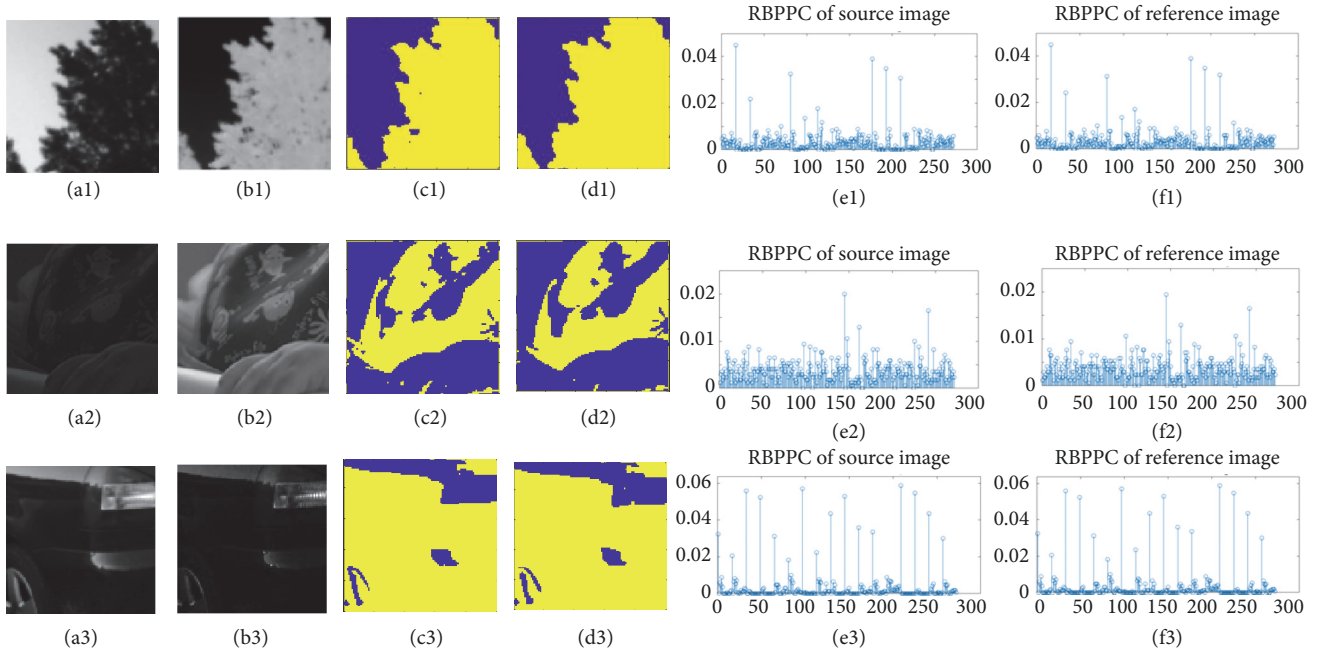


FIGURE 12: Examples of the multiple support region-based RBPPC descriptor: (a1), (b1) RGB/LWIR input image patch pair; (a2), (b2) an input image patch pair featuring different exposures; (a3), (b3) an input image patch pair featuring illumination changes; (c1), (c2), (c3) M_s from the respective source images; (d1), (d2), (d3) M_r from the respective reference images; (e1), (e2), (e3) multiple support region-based RBPPCs of the source images; (f1), (f2), (f3) multiple support region-based RBPPCs of the reference images.

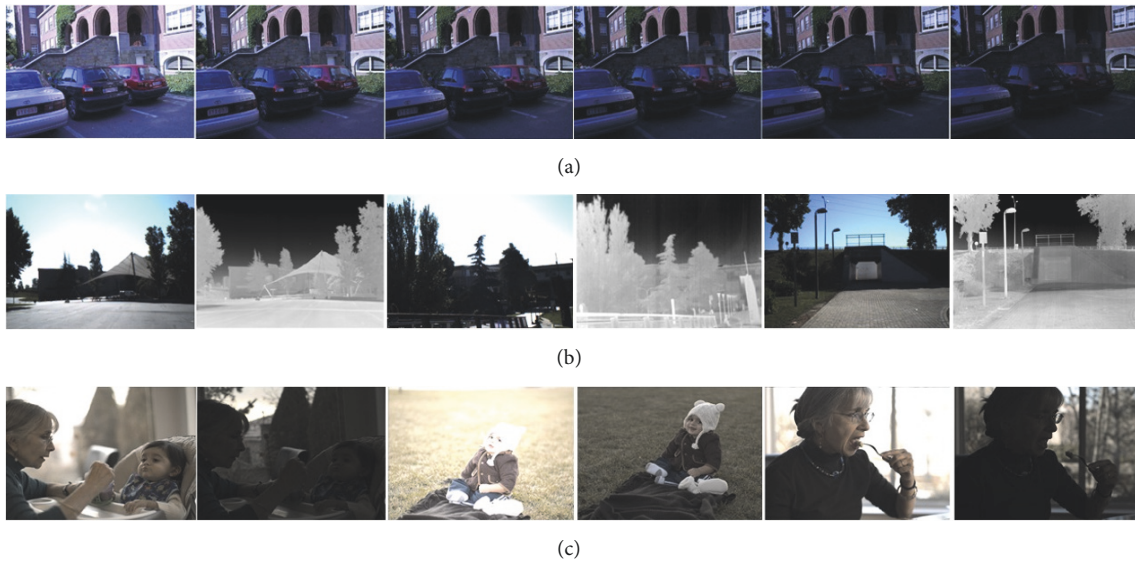


FIGURE 13: Examples of image matching test data: (a) illumination changes DB1, (b) multispectral DB3, (c) different exposures DB2.

image with the same scene), and three examples are shown as in Figure 13(b).

3.1. Parameters Analysis of the Proposed Feature Descriptor. Figure 14 shows the results of an extensive analysis of the RBPPC descriptor performance when varying its associated parameters, namely, the codebook patch scale a , the random sampling time n , and the threshold of the Chamfer distance D_{th} .

A successful image matching rate, P_{rate} , is employed as an evaluation indicator in order to evaluate its performance quantitatively.

$$P_{rate} = \frac{\text{number of successfully matched points}}{\text{total number of matching points}} \times 100\% \quad (12)$$

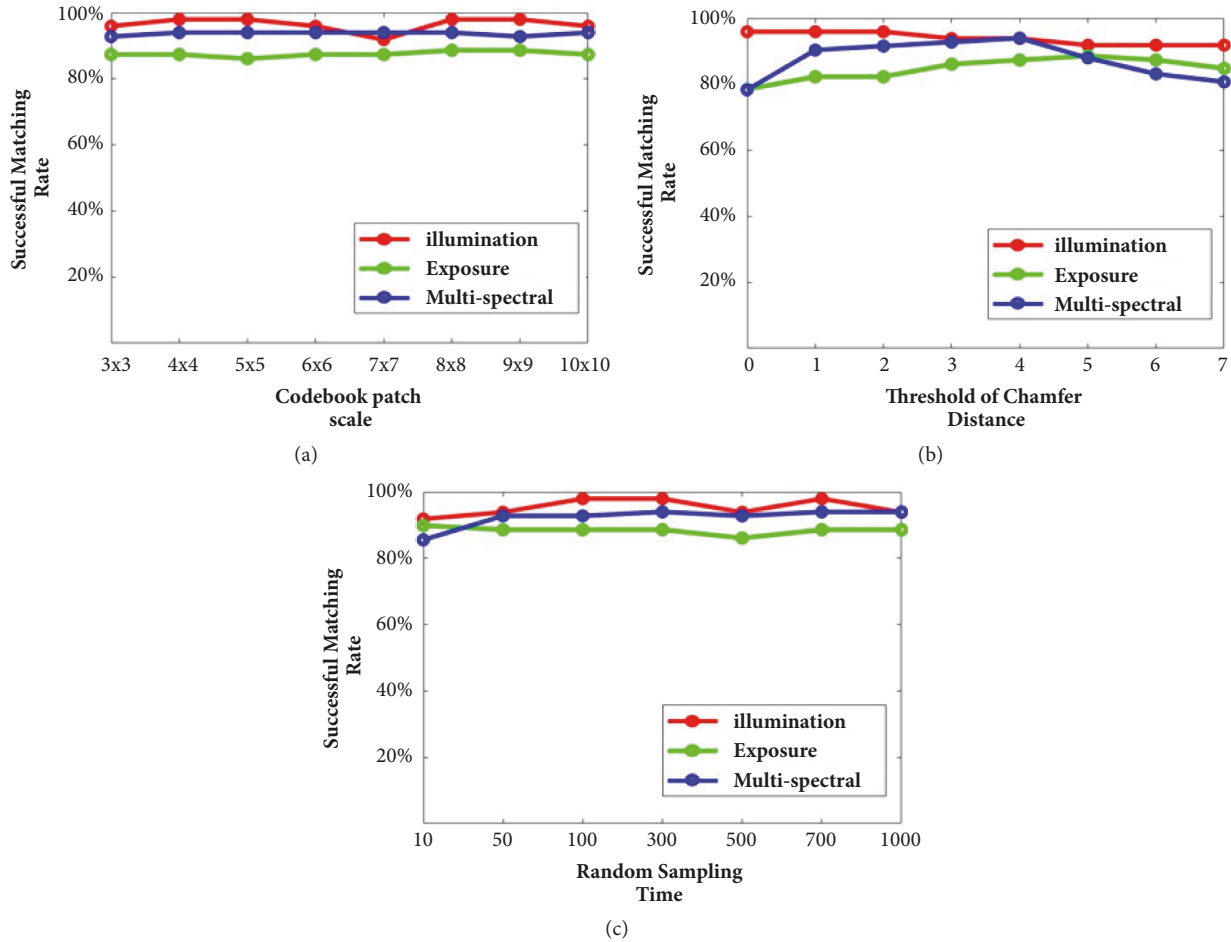


FIGURE 14: Image matching performance under varying parameters: (a) local patch size a , (b) threshold of the Chamfer distance D_{th} , (c) random sampling time n .

Note that, as Figure 14(a) shows, the matching performance of the proposed algorithm is not affected strongly by the local patch size a of a codebook with any of the test data types, since the performance curve is so flat. Varying the threshold of the Chamfer distance D_{th} has a larger effect on performance, as shown in Figure 14(b). With the illumination change data, increasing the threshold causes the matching quality to degrade. However, with the exposure change and the multispectral data, increasing the threshold across a certain range improves the matching quality. The reason for this is that the local patch clustering transform generates more similar responses under illumination changes than under other kinds of intensity changes. The dissimilar region is effectively removed by the large Chamfer distance threshold. Therefore, with a large Chamfer distance threshold, better matching quality is achieved under different exposures or multispectral images. Moreover, if the Chamfer distance is too large, the generated sampling candidate region will become too small and the proposed histogram feature will lose its discrimination; as a result, the matching quality will fall down. Likewise, as Figure 14(c) shows, the matching performance is also improved by increasing the random sampling time n within a certain range for all test data types.

3.2. Comparison Experiments. The image matching performances of the proposed RBPPC feature descriptor and other conventional descriptors, EOH [19], DASC [24], LBP [17], PCEOH [21], under intensity changes were compared. To evaluate the performance of the feature descriptors while avoiding bias due to the feature detector performance, we follow a similar approach to [33]. We extract the key-point descriptors from the source images and then project them into the corresponding reference image pair using different feature descriptor-based matching results. The ground truth homography information is provided by published databases [8, 31, 32] so that the ground truth and the performance of the different descriptors can be evaluated. The histogram intersection method [34] is employed as a matching measure for the feature descriptor, defined in

$$\text{sim}(f_1, f_2) = 1 - \left(\frac{\sum_{i=1}^{\text{dim}} \min(f_1, f_2)}{\sum_{i=1}^{\text{dim}} f_1} \right), \quad (13)$$

where f_1 and f_2 are the histogram feature descriptors and dim represents the dimensions of f_1 and f_2 .

The precision-recall curve [33] is selected as a common criterion that is used to evaluate local feature descriptors. The

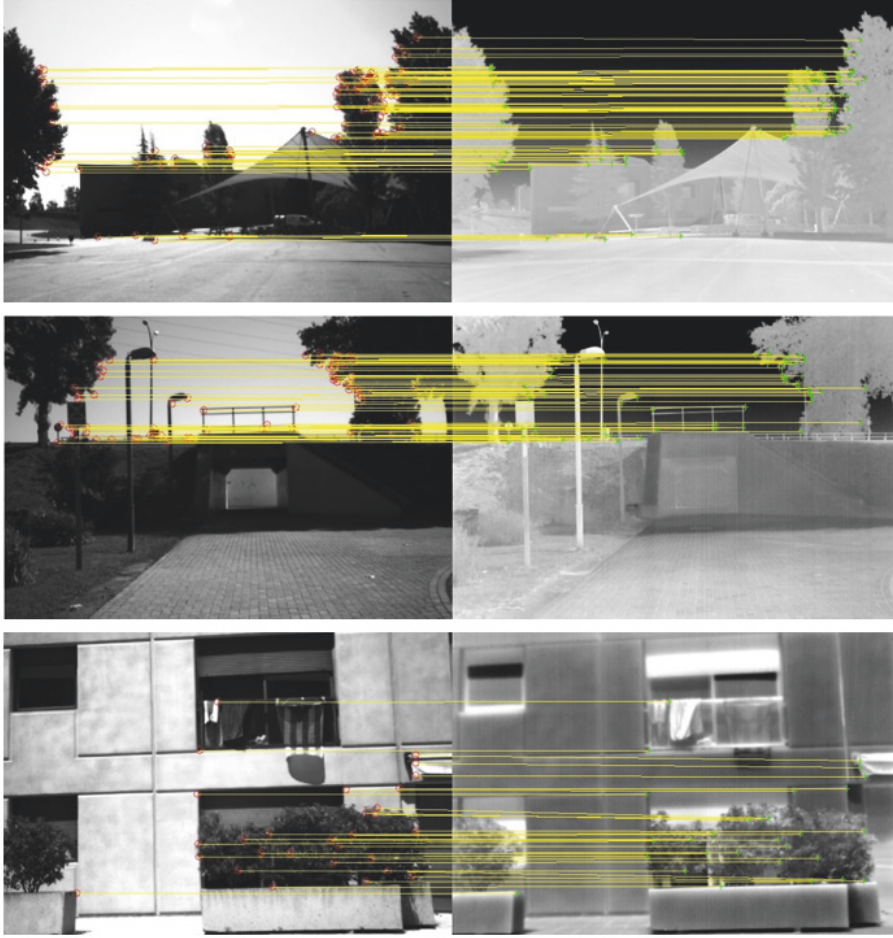


FIGURE 15: Key-point matching performance examples of test DB3 (multispectral images) with RBBPC feature.

curve is based on the number of correct matches and the total matches obtained for an image pair. “Recall” indicates the ratio of the correct matches to the corresponding number of the image pair. “Precision” denotes the ratio of the correct matches to the total matches. The “Recall” and “Precision” are defined in (14) and (15), respectively.

$$\text{recall} = \frac{\text{correct matches}}{\text{corresponding matches}} \times 100\%, \quad (14)$$

$$\text{1-precision} = \frac{\text{correct matches}}{\text{correct matches} + \text{flase matches}} \times 100\%. \quad (15)$$

Table 4 shows the image matching performances of the different feature descriptors with different test data types. It is clear that the performance of all the descriptors is better with the illumination change dataset than the other datasets. The reason for this is that the intensity changes under global illumination changes are approximately linear. However, the proposed RBPPC feature descriptor performs more effectively for all datasets than the other conventional methods, with high recall and low 1-precision.

Figures 15–17 show the key-point matching performances of proposed RBBPC feature descriptors under different test data types. Although there is a nonlinear intensity change between every test image pair (illumination change, different exposures, or multispectral images), key-point matching has been performed well with the help of the proposed RBBPC feature.

Figure 18 shows examples of the performances of the proposed RBPPC and other four conventional feature descriptors (EOH [19], LBP [17], DASC [24], and PCEOH [21]) under nonlinear intensity changes, that is, with RGB/LWIR multispectral image patch pairs.

4. Conclusions

A new feature descriptor is proposed in this paper for image matching under intensity changes. In contrast to conventional methods, a binary local patch clustering transform is proposed to transform the spaces of the original intensity change image pair. The values of the new spaces have higher similarities than the raw intensities, and a random binary pattern coding method converts the values of the transformed

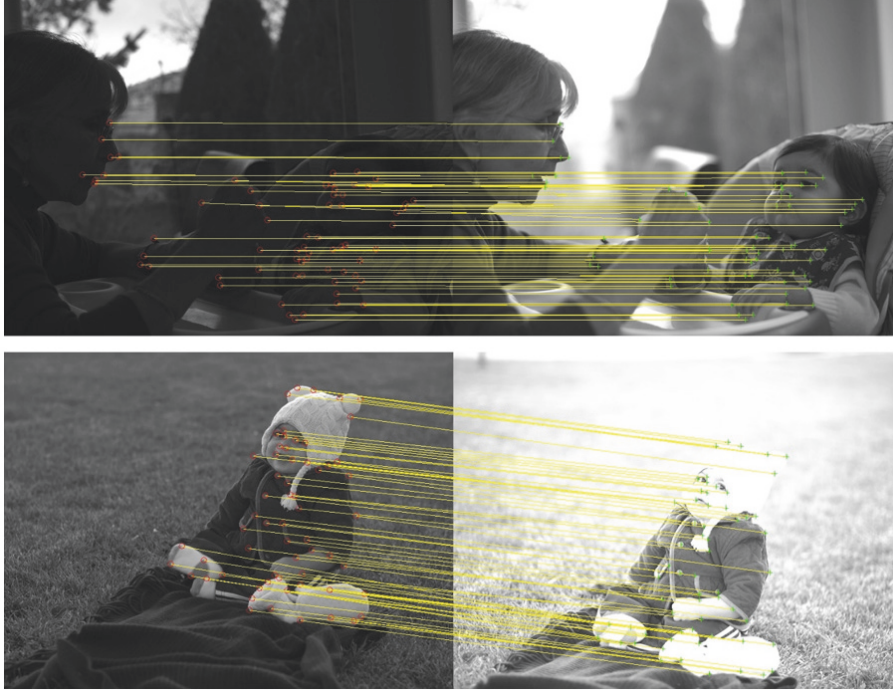


FIGURE 16: Key-point matching performance examples of test DB2 (different exposures) with RBPPC feature.

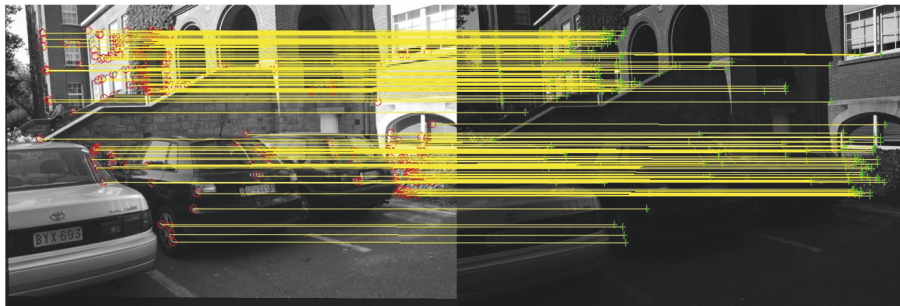


FIGURE 17: Key-point matching performances examples of test DB1 (illumination changes) with RBPPC feature.

spaces into feature histograms. Moreover, to enhance the discrimination of the proposed RBPPC, a Chamfer distance-based threshold is combined with a multiple spatial support region mask. The experimental results demonstrate that, given its high matching success rate, the proposed RBPPC descriptor performs more effectively than the conventional methods of image matching with intensity change image pairs.

Data Availability

The [DB1[1] image] data used to support the findings of this study are included within the article. The [DB2[2] image] data used to support the findings of this study are included within the article. The [DB3[3] image] data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Han Wang and Hanseok Ko came up with the main idea for the research; Han Wang performed the research and numerical analysis; Zhihuo Xu performed some additional parts of the experiments analysis and partially edited the paper; Hanseok Ko gave valuable suggestions on the revision of the paper and partially edited the paper.

Acknowledgments

This work was supported in part by the Natural Science Foundation of China under Grant 61872425, 61771265,

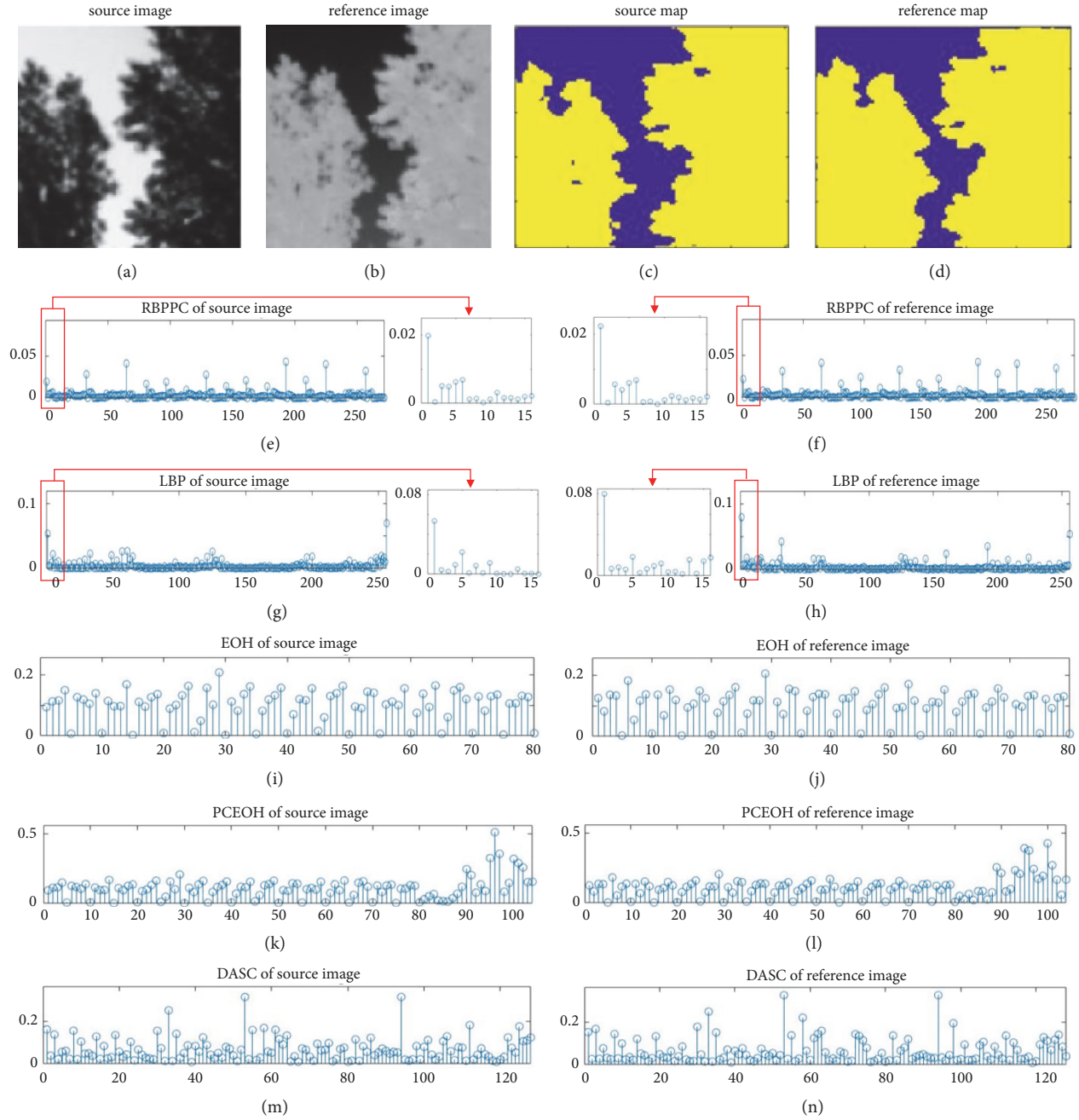


FIGURE 18: Example of different feature descriptors when processing the same RGB/LWIR image patch pair: (a) RGB source image, (b) LWIR reference image, (c) binary local patch transform of (a), (d) binary local patch transform of (b), (e) RBPPC descriptor of (a), (f) RBPPC descriptor of (b), (g) LBP descriptor of (a), (h) LBP descriptor of (b), (i) EOH descriptor of (a), (j) EOH descriptor of (b), (k) PCEOH descriptor of (a), (l) PCEOH descriptor of (b), (m) DASC descriptor of (a), (n) DASC descriptor of (b).

TABLE 4: Average 1-precision and recall for test DBs.

Test DB	DB1		DB2		DB3	
	Recall	1-Precision	Recall	1-Precision	Recall	1-Precision
EOH [19]	95%	22%	78%	45%	28%	82%
DASC [24]	96%	12%	83%	31%	36%	71%
LBP [17]	62%	54%	35%	67%	8%	90%
PCEOH [21]	91%	27%	85%	35%	24%	87%
Proposed	98%	9%	88%	24%	46%	63%

and 61801247, in part by the Nantong Science and Technology Bureau Foundation under Grant GY12016020 and GY12016017, in part by the University Science Research of Jiangsu Province Foundation under Grant 17KJB520029 and 17KJB510047, in part by Nantong University (17R31), and in part by the Basic Science Research Program of NRF (NRF-2017R1A2B4012720).

References

- [1] S. T. Li, H. T. Yin, and L. Y. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 12, pp. 3450–3459, 2012.
- [2] M. J. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Information Fusion*, vol. 27, pp. 198–214, 2015.
- [3] T. Simon, J. Valmadré, I. Matthews, and Y. Sheikh, "Kronecker-Markov Prior for Dynamic 3D Reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2201–2214, 2017.
- [4] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 7–42, 2002.
- [5] B. Kamolrat, W. A. C. Fernando, and M. Mrak, "3D motion estimation for depth information compression in 3D-TV applications," *IEEE Electronics Letters*, vol. 44, no. 21, pp. 1244–1245, 2008.
- [6] H. Wang, S. M. Yoon, D. K. Han, and H. Ko, "A feature descriptor based on the local patch clustering distribution for illumination-robust image matching," *Pattern Recognition Letters*, vol. 94, pp. 46–54, 2017.
- [7] Z. Wang, B. Fan, and F. Wu, "Local intensity order pattern for feature description," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 603–610, Barcelona, Spain, November 2011.
- [8] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust patch-based HDR reconstruction of dynamic scenes," *ACM Transactions on Graphics*, vol. 31, no. 6, 2012.
- [9] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," in *Proceedings of the Proceedings of ACM SIGGRAPH 2004*, pp. 664–672, August 2004.
- [10] D. Firmenichy, M. Brown, and S. Susstrunk, "Multispectral interest points for RGB-NIR image registration," in *Proceedings of the 2011 18th IEEE International Conference on Image Processing (ICIP 2011)*, pp. 181–184, Brussels, Belgium, September 2011.
- [11] M. Brown and S. Susstrunk, "Multi-spectral SIFT for scene category recognition," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 177–184, Colorado Springs, CO, USA, June 2011.
- [12] X. Shen, L. Xu, Q. Zhang, and J. Jia, in *Proceedings of the 13th Europe Conference on Compute Vision (ECCV)*, Multi-modal and multi-spectral registration for natural images, Ed., pp. 309–324, Zürich, Switzerland, 2014.
- [13] Y. Hacohen, E. Shechtman, and D. Lischinski, "Deblurring by Example Using Dense Correspondence," in *Proceedings of the 25th International Conference on Compute Vision (ICCV)*, pp. 2384–2391, Sydney, Australia, December 2013.
- [14] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [15] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425–436, 2009.
- [16] R. Gupta, H. Patil, and A. Mittal, "Robust order-based methods for feature description," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 334–341, June 2010.
- [17] F. Tang, S. H. Lim, N. L. Chang, and H. Tao, "A novel feature descriptor invariant to complex brightness changes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 2631–2638, IEEE, June 2009.
- [18] B. Kim, H. Yoo, and K. Sohn, "Exact order based feature descriptor for illumination robust image matching," *Pattern Recognition*, vol. 46, no. 12, pp. 3268–3278, 2013.
- [19] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, "Multispectral image feature points," *Sensors*, vol. 12, no. 9, pp. 12661–12672, 2012.
- [20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [21] T. Mouats, N. Aouf, A. D. Sappa, C. Aguilera, and R. Toledo, "Multispectral stereo odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1210–1224, 2015.
- [22] E. Shechtman and M. Irani, "Matching local self-similarities across images and videos," in *Proceedings of the 20th IEEE Conference on Compute Vision and Pattern Recognition (CVPR)*, pp. 17–22, Minneapolis, MN, USA, June 2007.
- [23] M. P. Heinrich, M. Jenkinson, M. Bhushan et al., "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Medical Image Analysis*, vol. 16, no. 7, pp. 1423–1435, 2012.
- [24] S. Kim, D. Min, B. Ham, M. N. Do, and K. Sohn, "DASC: Robust Dense Descriptor for Multi-Modal and Multi-Spectral Correspondence Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 9, pp. 1712–1729, 2017.
- [25] S. Kim, D. Min, B. Ham, M. N. Do, and K. Sohn, "DASC: Dense Adaptive Self-Correlation Descriptor for Multi-modal and Multispectral Correspondence," in *Proceedings of the 32th IEEE Conference on Compute Vision and Pattern Recognition (CVPR)*, pp. 2103–2112, Boston, MA, USA, June 2015.
- [26] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," *Expert Systems with Applications*, vol. 40, no. 1, pp. 200–210, 2013.
- [27] N. Nobuyuki, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [28] B. Wang, Y. Zhang, X. Wang, and C. Wu, "A new regularized minimum error thresholding method," *Transactions of Nanjing University of Aeronautics and Astronautics*, vol. 32, no. 4, pp. 355–364, 2015.
- [29] W. Niblack, *An Introduction to Digital Image Processing*, Prentice Hall, 1986.
- [30] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

- [31] <http://www.robots.ox.ac.uk/~vgg/research/affine/>.
- [32] C. A. Aguilera, A. D. Sappa, and R. Toledo, "LGHD: A feature descriptor for matching across non-linear intensity variations," in *Proceedings of the 22th International Conference on Image Processing (ICIP)*, pp. 178–181, Quebec City, Canada, September 2015.
- [33] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [34] K. Grauman and T. Darrell, "The pyramid match kernel: discriminative classification with sets of image features," in *Proceedings of the 17th International Conference on Compute Vision (ICCV)*, vol. 2, pp. 1458–1465, Beijing, China, October 2005.

