International Journal of Geo-Information

MDPI

Article

An Illumination Insensitive Descriptor Combining the CSLBP Features for Street View Images in Augmented Reality: Experimental Studies

Zejun Xiang [1,2], Ronghua Yang [1,*], Chang Deng [1], Mingxing Teng [1], Mengkun She [1] and Degui Teng [2]

1 School of Civil Engineering, Chongqing University, Chongqing 400045, China; xiangzj@cqkcy.com (Z.X.); chang_deng@cqu.edu.cn (C.D.); 20135772@cqu.edu.cn (M.T.); 20135774@cqu.edu.cn (M.S.)
2 Chongqing Survey Institute, Chongqing 401121, China; tengdg@cqkcy.com
* Correspondence: rh_yang@cqu.edu.cn; Tel.: +86-136-1825-6113

Received: 25 April 2020; Accepted: 31 May 2020; Published: 1 June 2020

Abstract: The common feature matching algorithms for street view images are sensitive to the illumination changes in augmented reality (AR), this may cause low accuracy of matching between street view images. This paper proposes a novel illumination insensitive feature descriptor by integrating the center-symmetric local binary pattern (CS-LBP) into a common feature description framework. This proposed descriptor can be used to improve the performance of eight commonly used feature-matching algorithms, e.g., SIFT, SURF, DAISY, BRISK, ORB, FREAK, KAZE, and AKAZE. We perform the experiments on five street view image sequences with different illumination changes. By comparing with the performance of eight original algorithms, the evaluation results show that our improved algorithms can improve the matching accuracy of street view images with changing illumination. Further, the time consumption only increases a little. Therefore, our combined descriptors are much more robust against light changes to satisfy the high precision requirement of augmented reality (AR) system.

Keywords: image feature matching; feature descriptor; CS-LBP; illumination robustness; street view images; augmented reality (AR)

1. Introduction

Augmented reality (AR) is an emerging form of experience in which the real world (RW) is enhanced by computer-generated content, which is tied to specific locations and/or activities. In simple terms, AR allows digital content to be seamlessly overlaid and mixed into our perceptions of the real world [1]. AR technology can be divided into two categories: marker-based AR and markerless-based AR. Marker-based AR is to use visual features or an object to be a trigger, while markerless-based AR is to use some technology to detect the relative position (feature matching) between virtual objects and the real world [2]. The principle of markerless-based AR is to extract feature points from template objects through a series of algorithms, such as SURF, ORB, FERN, etc., and record or learn these feature points. The tracking based on image processing uses natural features that, color, shape, texture and interest point, in images to calculate a camera's pose [3]. It uses the homography matrix between adjacent frame images obtained by image matching to solve the position and pose of the camera for the registration [4]. The vision system generates in a pre-processing phase with a huge quantity of 3D-features. GIS (Geographic Information System) technological researches are carried out to design and build a geographic database that can store and query these and can add the features that are generated by the augmented reality system [5].

Street view image is the main component of a real scene in AR system, since the illumination changes and the different photosensitivity of different cameras will cause the complex brightness changes of the same street view scene after imaging, which makes it difficult to match the different images [6]. Aiming at the difficulties of image feature matching caused by illumination changes, researchers have proposed many related image feature matching algorithms. Lowe proposed a scale-invariant local feature detection and description algorithm called SIFT [7]. Thanks to its high uniqueness and robustness, SIFT has been demonstrated as the state-of-the-art algorithm in computer vision. However, due to the loss of information and the changes of brightness of the decolorized images are not considered, the high dimensional descriptor makes it very slow when extracting information from features. To speed it up, Bay proposed the SURF algorithm for the same purpose with lower computational complexity, but illumination changes remain problematic [8]. Tola proposed another local feature description algorithm on the basis of SIFT, which convolves several Gaussian filter functions in different scales with several directional diagrams of the original image [9]. The efficiency of the algorithm has improved, but the illumination changes is not considered [10]. Leutenegger proposed a feature detector and descriptor called BRISK [11], which uses the FAST detector [12] and a binary feature descriptor, thus is faster than SURF and performs well regarding rotation and scale, while not robust against illumination changes [6]. Developed from the oriented FAST detector and rotated BRIEF descriptor [12], Rublee [13] proposed a fast feature detection and description algorithm called ORB, while it is also sensitive to illumination changes [6]. Following this, Alahi proposed FREAK to improve BRISK [14], which adopts a better sampling pattern that is very similar to the retinal ganglion cells of human eyes when receiving visual information. It is faster but less precise than BRISK. Alcantarilla proposed a feature detection and description algorithm that operates in a nonlinear scale space [15]. Through using additive operator splitting (AOS) techniques [16] and variable conductance diffusion to construct the nonlinear scale space with arbitrary step size, KAZE solves the problem of blurred boundaries and losing details of the image caused by Gaussian scale space like SIFT and SURF. Nevertheless, the KAZE algorithm improves the localization accuracy, but the computational efficiency becomes lower. Later, Alcantarilla proposed a more accurate and efficient feature detection and description algorithm called AKAZE [17], which introduces the fast explicit diffusion (FED) scheme to solve the partial differential equations, which is much faster and more accurate than AOS [18]. Moreover, a highly efficient modified-local difference binary (M-LDB) descriptor is introduced to increase the robustness against rotation and scale invariance, while it gets low precision in the scene of clear lighting changes. Gevrekci proposed an illumination robust interest point detection algorithm, but its high complexity makes its real-time performance extremely poor [19]. Liu improved the Hu invariant moment to increase its illumination robustness, while this could cause the loss of image information, leading to a relatively low precision [20]. Ouyang matches images with different brightness in affine space and its performance is good, but the large amount of computation makes it not suitable for real-time image matching [21]. Lyu uses the non-linear correlation of grayscale information in high-dimensional vector space to improve the illumination robustness of image matching, but it will cause the loss of image information as well [22].

To solve the above problems, this paper integrates the center-symmetric local binary pattern (CS-LBP), a simple but strong illumination-invariant feature, into eight common feature correspondence algorithms [23]. We conduct experiments on the combined descriptors, the CSLBP and eight original descriptors in five different illumination scenes. The experimental results show that the improved algorithms can increase the precision of image matching between street view images with different illumination changes. In addition, they can as well satisfy the real-time requirements in AR systems.

2. Methods

2.1. CS-LBP Feature Descriptor

The CS-LBP [24] is a simplified local binary pattern (LBP) [25], which can be used to describe the relationship of the pixel intensity between the point and its local region. The implementation steps of CS-LBP descriptor can be summarized as follows:

Step-1: Suppose $P_i(x, y, \sigma, \theta)$ is a keypoint identified by a detector, where (x, y) represent the pixel coordinate in the original image, σ defines the scale and θ denotes the orientation of the keypoint.

Step-2: Take a 9×9 area centered around the keypoint (inside the black-line in Figure 1a) at the scale space of σ, which is used compute the CS-LBP feature descriptor of the keypoint (here we expand a pixel outward to compute the CS-LBP feature of the pixel at the edge). Suppose j is an arbitrary pixel within the 9×9 area, as shown in Figure 1a. Then, rotate this region to the orientation of the keypoint at an angle of θ to obtain the rotation invariance, as shown in Figure 1b.

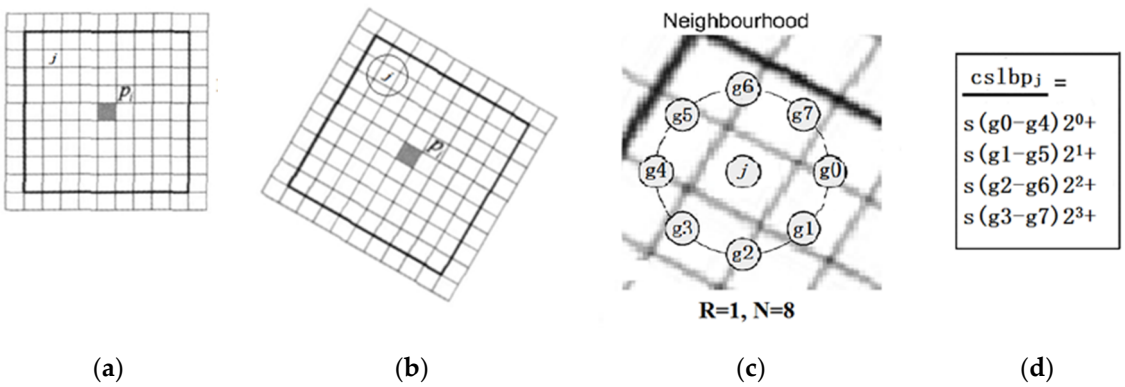


Figure 1. Steps to compute center-symmetric local binary pattern (CS-LBP) features. (**a**) Area to compute the CS-LBP features centered around the keypoint; (**b**) the area after rotation; (**c**) CS-LBP feature for a neighborhood of 8 pixels of j, when $R = 1$ and $N = 8$; (**d**) binary pattern for CS-LBP feature of j.

Step-3: Compute the CS-LBP feature of the pixel j according to Equation (1) as $cslbp_j$, with $j = 1, 2, \ldots, 81$.

$$cslbp_j = CS - LBP_{R,N,T}(x, y) = \sum_{i=0}^{N/2-1} s(g_i - g_{i+N/2})2^i,\ s(k) = \begin{cases} 1 & k > T \\ 0 & otherwise \end{cases}, \tag{1}$$

where g_i and $g_{i+N/2}$ represent the normalized gray values of a center-symmetric pairs of pixels of N equally spaced pixels on a circle with a radius of R, and T defines the threshold of normalized gray value. In this paper we take $R = 1$, $N = 8$, $T = 0.01$, see Figure 1c.

Step-4: If the keypoint is detected by BRISK, ORB, FREAK or AKAZE algorithm (where the original descriptors of these four algorithm is binary), we construct an 81-dimensional binary CS-LBP description vector $\boldsymbol{V_i}$ for it by 81 $cslbp_j$ according to Equation (2). Instead, if it is detected by SIFT, SURF, DAISY or KAZE algorithm, we need to further normalize the binary vector $\boldsymbol{V_i}$ by Equation (3) as a non-binary vector $\boldsymbol{U_i}$ [26].

$$\boldsymbol{V_i} = [cslbp_1\ cslbp_2\ \cdots\ cslbp_{81}]. \tag{2}$$

$$\frac{\boldsymbol{V_i}}{\|\boldsymbol{V_i}\|} \rightarrow \boldsymbol{U_i} \tag{3}$$

where $\|\boldsymbol{V_i}\|$ represents the module of $\boldsymbol{V_i}$.

Figure 1 shows the processing of computing CS-LBP feature, and Figure 2 shows the flow chart of constructing CS-LBP description vector.

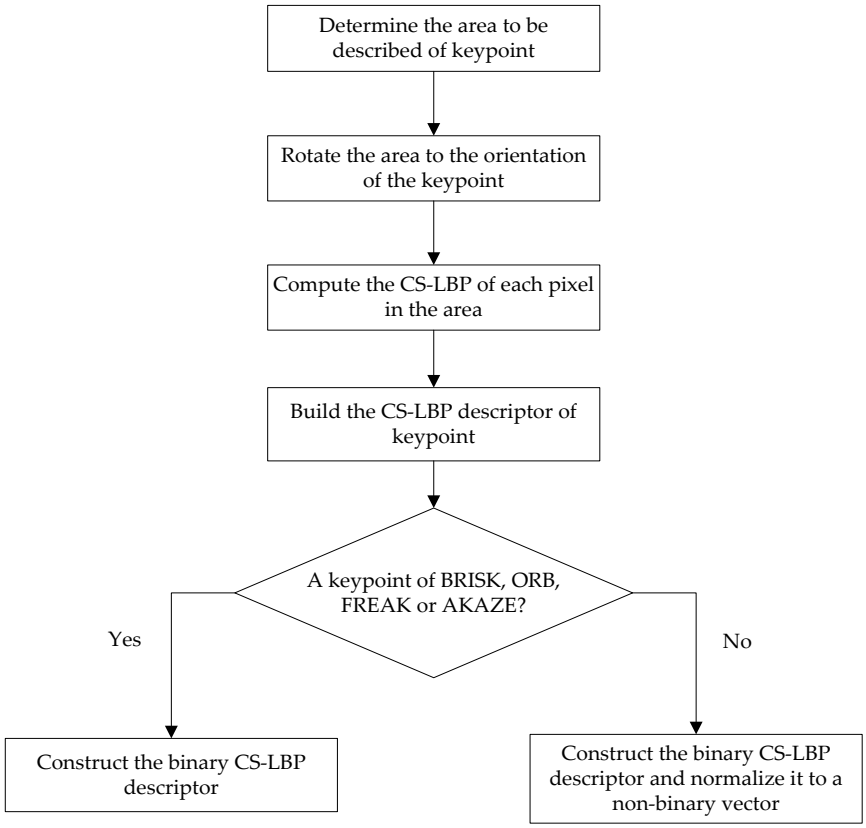


Figure 2. Flow chart for the CS-LBP description vector.

2.2. Combined Descriptors

Firstly, we obtain the keypoints and their original descriptors through eight original image feature matching algorithms (SIFT, SURF, DAISY, BRISK, ORB, FREAK, KAZE, AKAZE). Then we compute the CS-LBP descriptor of each keypoint. Afterwards, we add the CS-LBP descriptors to eight original descriptors, respectively, to construct eight combined descriptors, they are denoted correspondingly as SIFT-CSLBP, SURF-CSLBP, DASIY-CSLBP, BRISK-CSLBP, ORB-CSLBP, FREAK-CSLBP, KAZE-CSLBP and AKAZE-CSLBP. Lastly, we match the keypoints of two images by the nearest distance ratio method. The implementation steps are summarized as follows:

Step-1: The keypoints sets of the two images are detected by one of the eight original feature detection algorithms.

Step-2: The original descriptor of each keypoint is computed by one of the eight original feature description algorithms.

Step-3: The CS-LBP descriptor of each keypoint is computed by the steps in Section 2.1.

Step-4: The original descriptor is simply concatenated with the CS-LBP descriptor, and then we can get the new combined descriptor with increased dimension.

Step-5: Match the keypoints in the querying image and the training image by the nearest distance ratio method (the ratio here is 0.8 (Lowe, 2004)), e.g., Euclidean distance for non-binary descriptor (SIFT-CSLBP, SURF-CSLBP, DAISY-CSLBP or KAZE-CSLBP), while Hamming distance for binary descriptor (BRISK-CSLBP, ORB-CSLBP, FREAK-CSLBP or AKAZE-CSLBP).

The flow chart for calculating the combined descriptors and matching the keypoints is shown in Figure 3. The original feature detection algorithm in the chart can be one of the mentioned eight feature detection and matching algorithms.

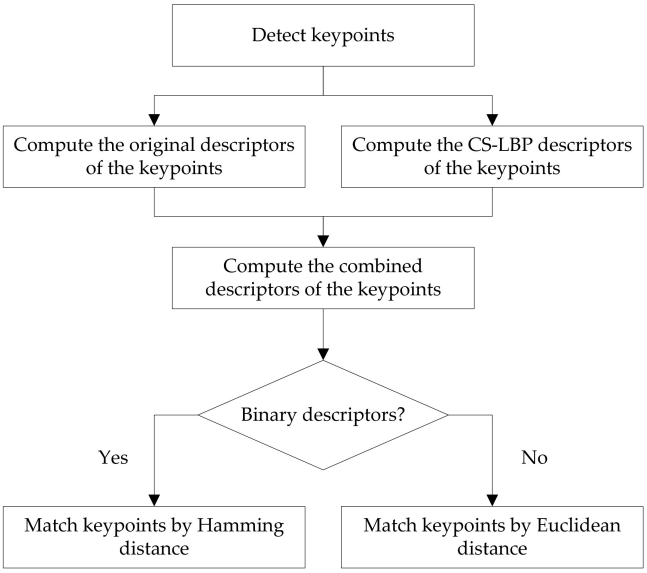


Figure 3. Flow chart for calculating the combined descriptors and matching the keypoints.

3. Experiments and Results

We first introduce the data and method of the experiment as well as the evaluation indices for the algorithm performance. Then we design a series of experiments to compare the matching performance between the combined descriptors, the CS-LBP descriptors and the original descriptors. We finally design experiments to compare the computational and matching speed of the three types of descriptors.

3.1. Datasets

The matching experiments were performed on five street view image sequences provided by Mikolajczyk, which are the benchmark sequences used to test the performance of image matching algorithms [27]. The specific presentation of each sequence is as follows:

- Leuven sequence

Leuven sequence only has illumination changes, and it contains six images (Figure 4). Each image has a size of 921 × 614 pixels.

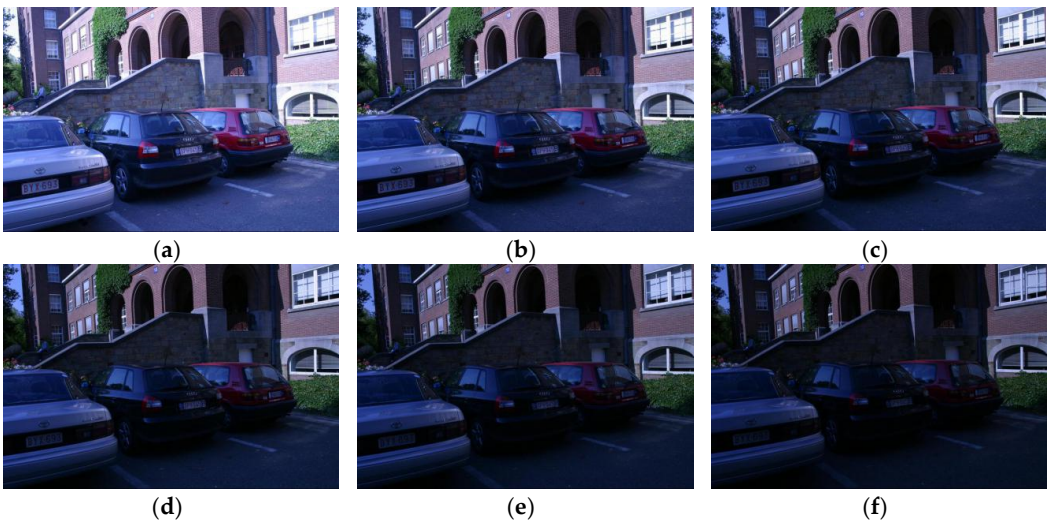


Figure 4. Leuven sequence. (**a**) Leuven 1; (**b**) Leuven 2; (**c**) Leuven 3; (**d**) Leuven 4; (**e**) Leuven 5; and (**f**) Leuven 6.

- Bikes sequence

Bikes sequence has blur changes. The Normal Bikes (Figure 5a) and New Bikes (Figure 5b) with increasing blur are selected for the experiment. The exposures of New Bikes are adjusted by Photoshop software to simulate changing illumination, which are −1, −2, −3, +1, +2, +3, respectively. The magnitude of illumination changes increases with the absolute value of the exposure. Specifically, negative exposure means underexposure, indicating that the image appears darker, while positive means overexposure, which suggests a brighter image. After processing, this sequence has both blur and illumination changes. Each image has a size of 1000 × 700 pixels (Figure 5).

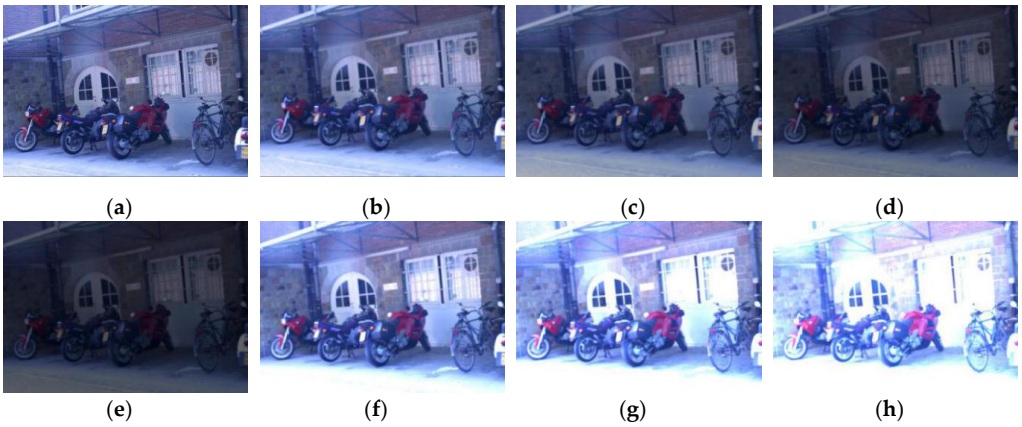


Figure 5. Bikes sequence. (**a**) Normal Bikes; (**b**) New Bikes with increasing blur; (**c**) New Bikes in an underexposure of −1; (**d**) New Bikes in an underexposure of −2; (**e**) New Bikes in an underexposure of −3; (**f**) New Bikes in an overexposure of +1; (**g**) New Bikes in an overexposure of +2; and (**h**) New Bikes in an overexposure of +3.

- Boat sequence

Boat sequence has scale and rotation changes. The Normal Boat (Figure 6a) and New Boat (Figure 6b) with a scale reduction factor of 1.9 and a rotation angle of 45 degrees are selected for experiment. The exposures of New Boat are also adjusted by Photoshop software as Bikes sequence. After that, this sequence has scale, rotation and illumination changes. Each image has a size of 800 × 640 pixels (Figure 6).

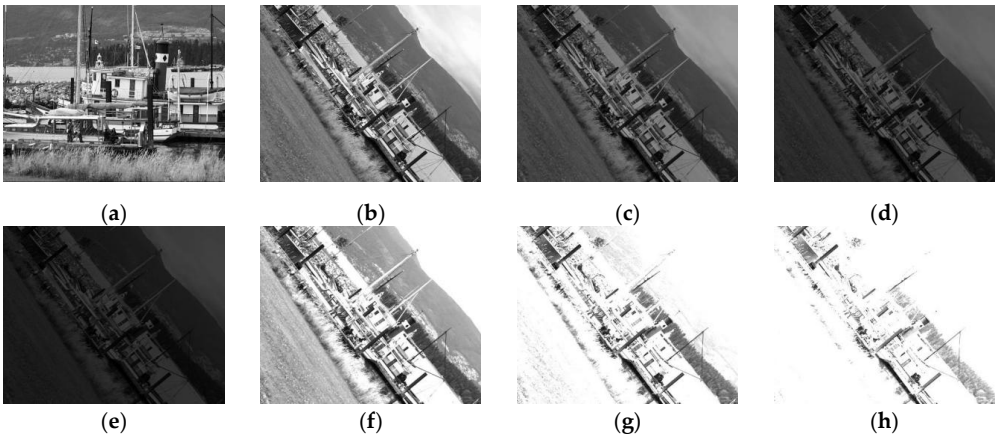


Figure 6. Boat sequence. (**a**) Normal Boat; (**b**) New Boat with a scale reduction factor of 1.9 and a rotation angle of 45 degrees; (**c**) New Boat in an underexposure of −1; (**d**) New Boat in an underexposure of −2; (**e**) New Boat in an underexposure of −3; (**f**) New Boat in an overexposure of +1; (**g**) New Boat in an overexposure of +2; (**h**) New Boat in an overexposure of +3.

- Graffiti sequence

Graffiti sequence has viewpoint changes. The Normal Graffiti (Figure 7a) and New Graffiti (Figure 7b) with a viewpoint changing angle of 40 degrees are selected. The same as the Bikes sequence, the exposures of New Graffiti are adjusted by Photoshop software. After that, this sequence has both viewpoint and illumination changes. Each image has a size of 800 × 640 pixels (Figure 7).

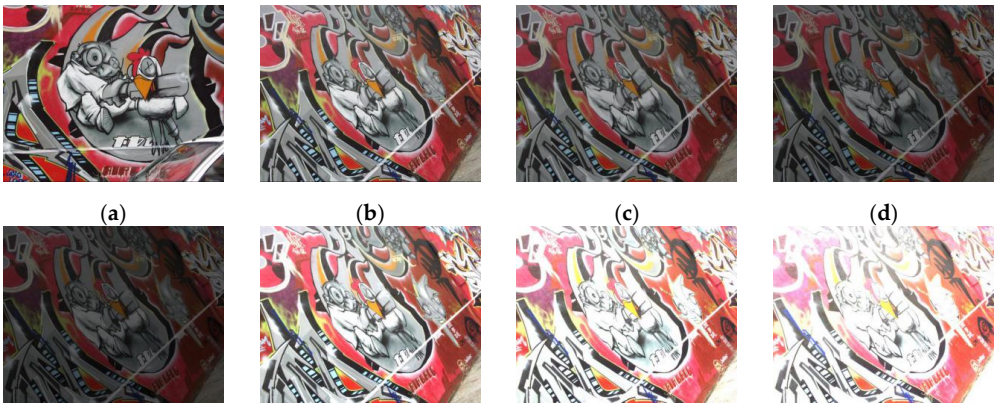


Figure 7. Graffiti sequence. (**a**) Normal Graffiti; (**b**) New Graffiti with a viewpoint changing angle of 40 degrees; (**c**) New Graffiti in an underexposure of −1; (**d**) New Graffiti in an underexposure of −2; (**e**) New Graffiti in an underexposure of −3; (**f**) New Graffiti in an overexposure of +1; (**g**) New Graffiti in an overexposure of +2; (**h**) New Graffiti in an overexposure of +3.

- Ubc sequence

The Ubc sequence has JPEG compression changes. The Normal Ubc (Figure 8a) and New Ubc (Figure 8b) with a JPEG compression of 90% are selected for the experiment. Similarly, the exposures of New Ubc are adjusted by Photoshop software as Bikes sequence. After that, this sequence has both JPEG compression and illumination changes. Each image has a size of 800 × 640 pixels (Figure 8).

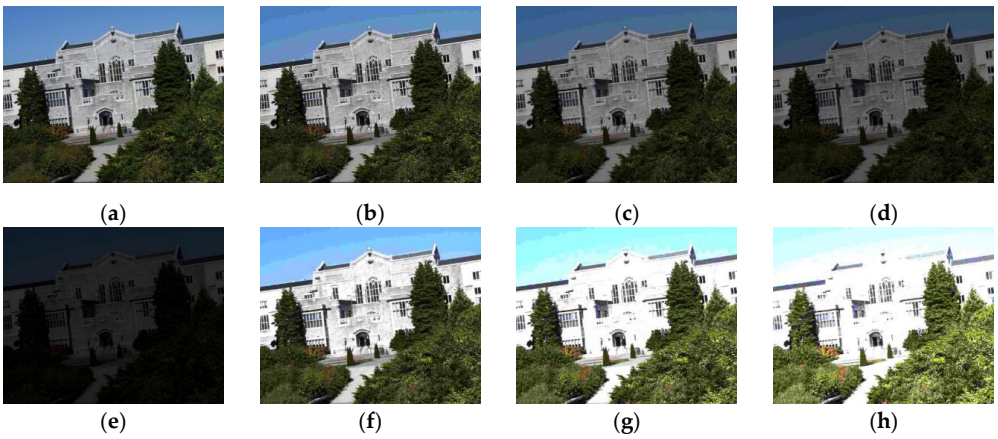


Figure 8. Ubc sequence. (**a**) Normal Ubc; (**b**) New Ubc with a JPEG compression of 90%; (**c**) New Ubc in an underexposure of −1; (**d**) New Ubc in an underexposure of −2; (**e**) New Ubc in an underexposure of −3; (**f**) New Ubc in an overexposure of +1; (**g**) New Ubc in an overexposure of +2; (**h**) New Ubc in an overexposure of +3.

3.2. Experiment Method

In order to compare the matching performance and the speed of the combined descriptors (SIFT-CSLBP, SURF-CSLBP, DAISY-CSLBP, BRISK-CSLBP, ORB-CSLBP, FREAK-CSLBP, KAZE-CSLBP,

AKAZE-CSLBP) and the CSLBP descriptors with eight original descriptors (SIFT, SURF, DAISY, BRISK, ORB, FREAK, KAZE and AKAZE) (where the CSLBP algorithm is the method that we replace the feature descriptor in the original algorithms by the CSLBP feature descriptor, and the only difference between these two algorithms is the different description), this paper tests five sequences in Section 3.1 in a programming environment of Visual Studio 2013 VC++ and OpenCV 3.3. The specific experimental processes are as follows:

Step-1: Selection of evaluation indices for the algorithm performance.

The performance of image feature matching algorithm commonly includes the calculation accuracy and efficiency in the detection, description and matching of keypoints [28]. Thus, a good algorithm should get more correct matches and lower time consumption. Therefore, we select the recall, precision and operating time [28] to evaluate the performance of our improved algorithms.

Step-2: Comparison of performance of the original and the improved algorithms.

We use the subfigure (a) of Figures 4–8 to match against the others within the sequence by using the original, the CSLBP and the combined descriptors, respectively.

Step-3: Comparison of speed of the original and the improved algorithms.

Since the feature detection step in each improved algorithm is identical to that of the original algorithm, we only compare the computational and matching speed of each description vector. We chose the Figure 1a of Leuven sequence for testing. Firstly, we extract 30 keypoints to construct the original, the CSLBP and the improved descriptors, respectively. Then, we match the 30 keypoints against the same 30 keypoints by using the original, the CSLBP and the improved descriptors, respectively. The time consumption for each descriptor of constructing and matching for 30 keypoints are recorded.

3.3. Evaluation Indices

The recall is computed as the ratio between the number of correctly matched keypoints and the number of corresponding keypoints [27] as shown in Equation (4). The precision is represented by the number of correct matches with respect to the number of total matches as shown in Equation (5).

$$recall = \frac{\#correct\ matches}{\#correspondence}, \tag{4}$$

$$precision = \frac{\#correct\ matches}{\#total\ matches}, \tag{5}$$

where #*correct matches* is the number of correct matches, #*correspondence* is the number of corresponding keypoints which represents the total number of possible correct matches including the matched and unmatched keypoints, while #*total matches* is the number of total actual matches including the correct and false matches.

To verify the correct matches, we use the criterion proposed by Mikolajczyk [28]. The match between two keypoints is correct if the pixel error in relative location is less than 3 pixels:

$$\left\|(x_a, y_a) - H\bullet(x_b, y_b)\right\| < 3, \tag{6}$$

where (x_a, y_a) and (x_b, y_b) are the pixel coordinate of the keypoint in training and querying image, respectively, H is the homography between two images.

3.4. Influences of Only Illumination Changes

With respect to Leuven sequence (Figure 4), we use Figure 4a to match against other images within the group with the original, the CSLBP and the combined descriptors, respectively. The recall and precision in the matching results of each descriptors are shown in Figures 9 and 10. The red line represents for the original descriptors, green line for the CSLBP descriptors and blue line for the combined descriptors. The reduced values of recall for the CSLBP and combined descriptors compared to the original descriptors are computed in Table 1, of which in bold is the highest decrease of this

descriptor, while the improved values of precision are in Table 2, of which in bold is the highest increase of this descriptor. From those figures and tables, we can get the following conclusions:

1. In most scenes, the recall of the CSLBP descriptor is far lower than that of the original descriptors, while the combined descriptors are just slightly lower.
2. The precision of the CSLBP descriptors is sometimes higher than that of the original descriptors and sometimes lower, while the combined descriptors are always higher.
3. Among the combined descriptors, the FREAK-CSLBP descriptor demonstrates the best matching performance with a less reduced recall (within 13.8%) and a higher precision (over 90%).

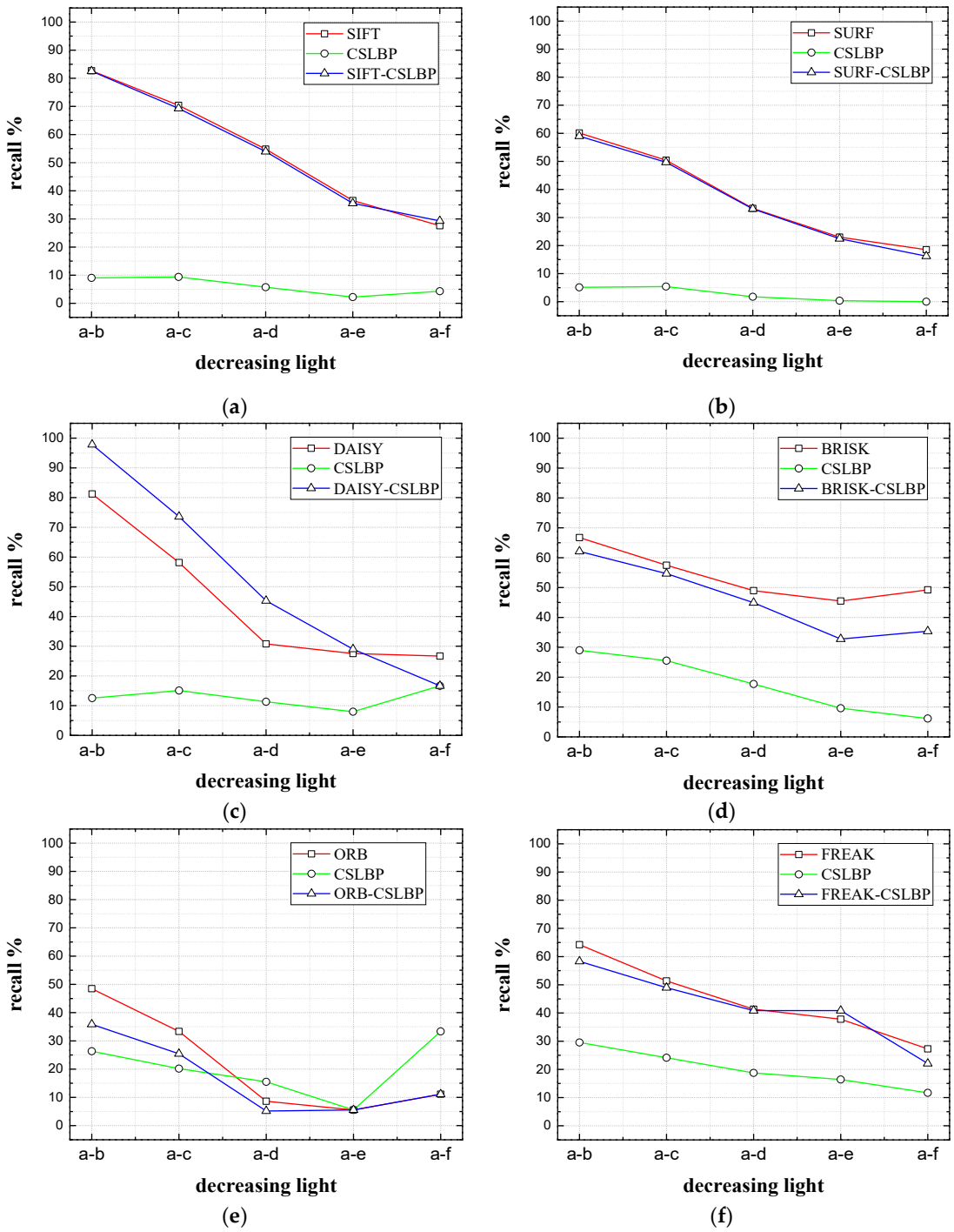


Figure 9. *Cont.*

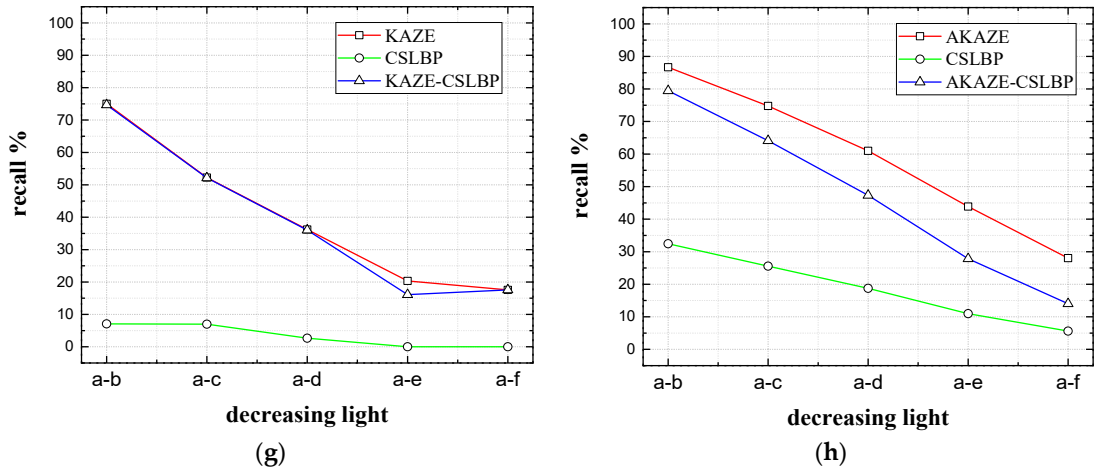


Figure 9. Matching results of the recall of Leuven sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Leuven sequence, and the Y label denotes the recall for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

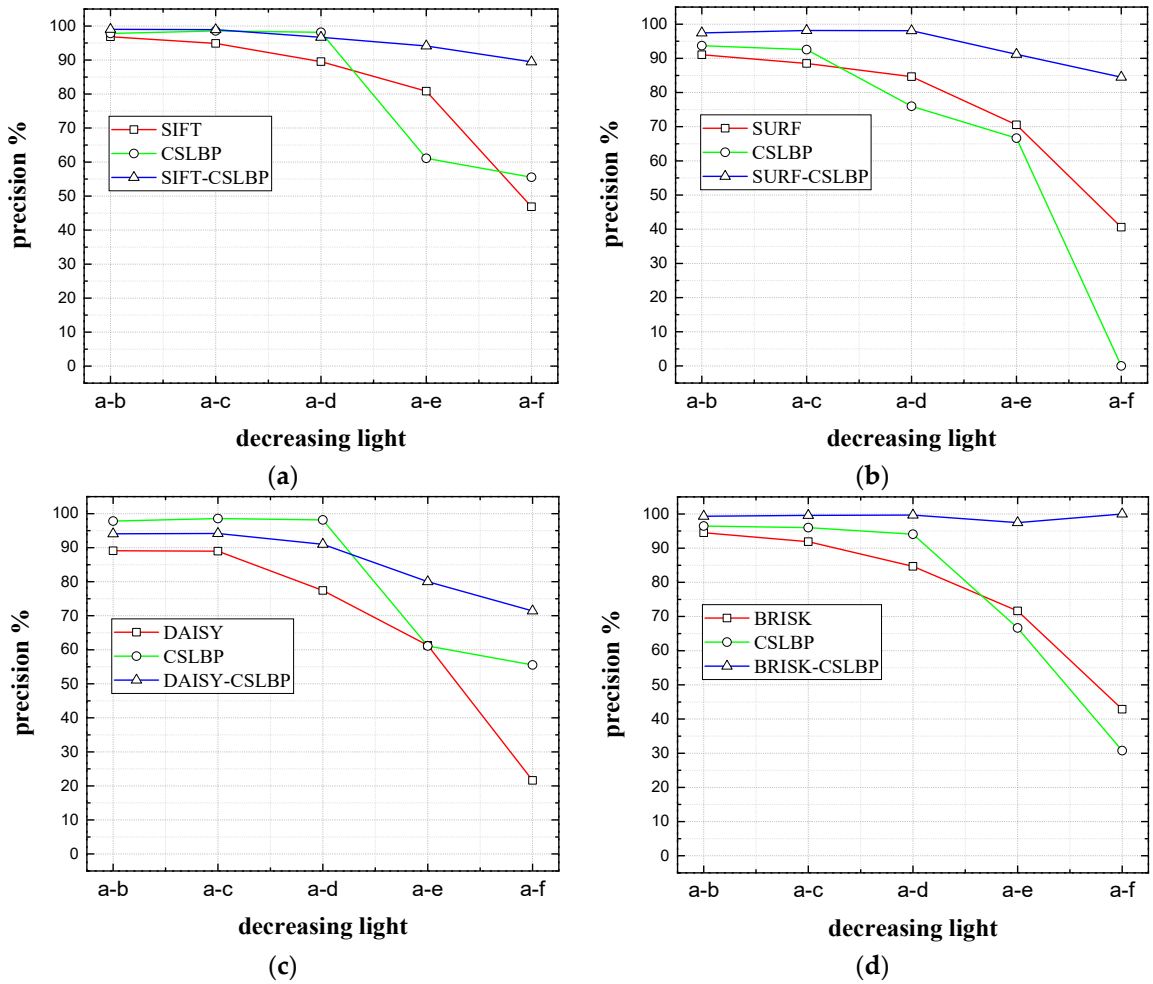


Figure 10. *Cont.*

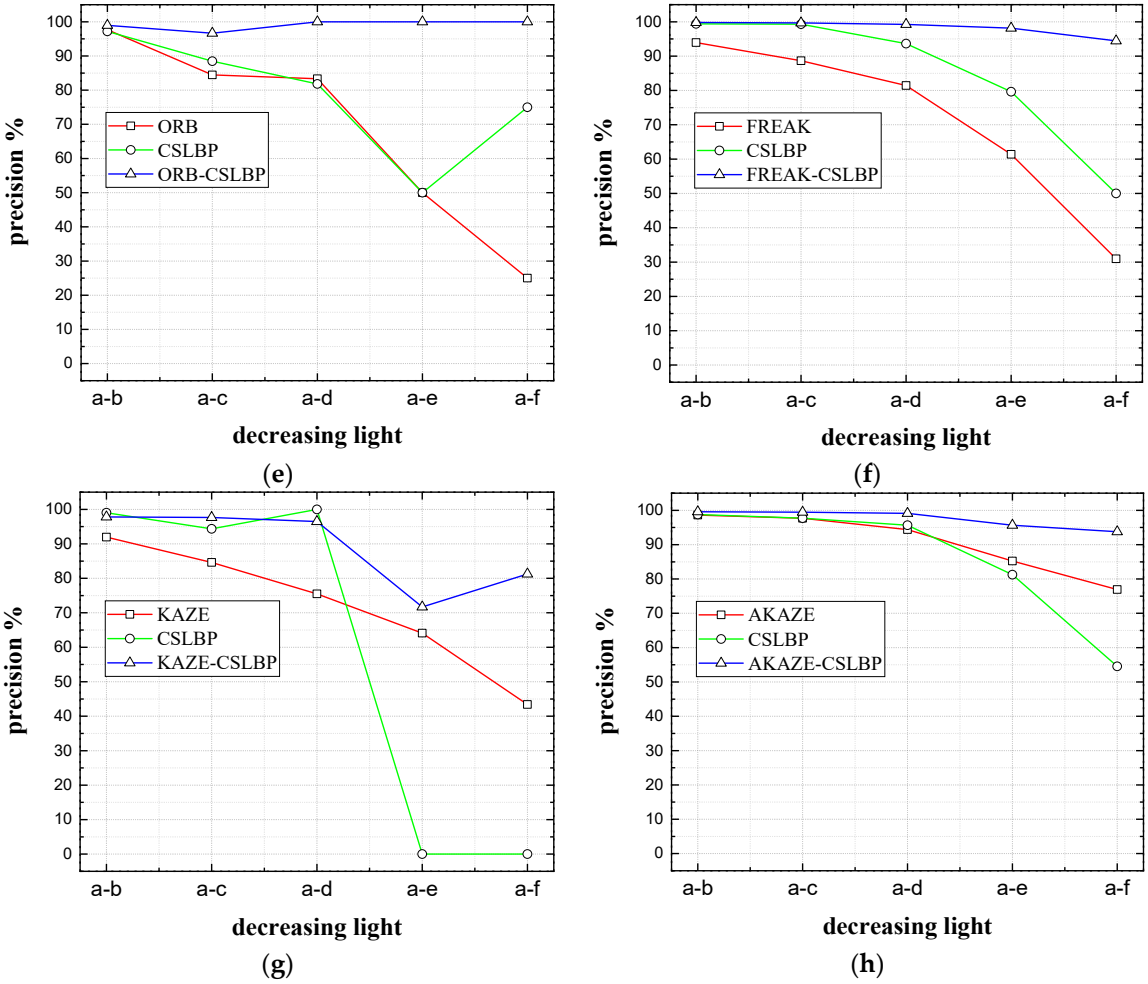


Figure 10. Matching results of the precision of Leuven sequence for the eight original, the CSLBP, and the improved descriptors. The X label represents the matched images of Leuven sequence, and the Y label denotes the precision for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; and (**h**) AKAZE features.

Table 1. The reduced values of recall for the CSLBP and the combined descriptors compared to the original descriptors of Leuven sequence.

Features	Descriptors	a-b	a-c	a-d	a-e	a-f
SIFT	CSLBP	**−73.6%**	−61.0%	−49.1%	−34.3%	−23.2%
	SIFT-CSLBP	−0.2%	**−1.1%**	−0.9%	−1.0%	1.7%
SURF	CSLBP	**−55.0%**	−45.0%	−31.5%	−22.6%	−18.5%
	SURF-CSLBP	−1.1%	−0.7%	−0.2%	−0.5%	**−2.3%**
DAISY	CSLBP	**−68.7%**	−43.1%	−19.4%	−19.6%	−10.0%
	DAISY-CSLBP	16.6%	15.4%	14.5%	1.4%	**−10.0%**
BRISK	CSLBP	−37.7%	−31.9%	−31.2%	−35.9%	**−43.1%**
	BRISK-CSLBP	−4.6%	−2.8%	−4.0%	−12.7%	**−13.8%**
ORB	CSLBP	**−22.1%**	−13.1%	6.9%	0	22.2%
	ORB-CSLBP	**−12.6%**	−7.9%	−3.4%	0	0
FREAK	CSLBP	**−34.7%**	−27.2%	−22.6%	−21.4%	−15.6%
	FREAK-CSLBP	**−5.9%**	−2.4%	−0.5%	3.1%	−5.2%
KAZE	CSLBP	**−67.9%**	−45.3%	−33.6%	−20.3%	−17.6%
	KAZE-CSLBP	−0.3%	−0.1%	−0.1%	**−4.2%**	0
AKAZE	CSLBP	**−54.2%**	−49.2%	−42.2%	−32.9%	−22.4%
	AKAZE-CSLBP	−7.2%	−10.7%	−13.6%	**−16.0%**	−14.0%

Table 2. The improved values of precision for the CSLBP and the combined descriptors compared to the original descriptors of Leuven sequence.

Features	Descriptors	a-b	a-c	a-d	a-e	a-f
SIFT	CSLBP	1.0%	3.7%	8.6%	−19.7%	**8.7%**
	SIFT-CSLBP	2.2%	4.0%	7.2%	13.3%	**42.6%**
SURF	CSLBP	2.6%	**4.1%**	−8.6%	−3.9%	−40.6%
	SURF-CSLBP	6.4%	9.7%	13.4%	20.6%	**43.9%**
DAISY	CSLBP	8.7%	9.6%	20.7%	−0.2%	**33.9%**
	DAISY-CSLBP	5.0%	5.2%	13.6%	18.7%	**49.8%**
BRISK	CSLBP	2.0%	4.1%	**9.4%**	−5.0%	−12.1%
	BRISK-CSLBP	4.8%	7.7%	15.0%	25.8%	**57.1%**
ORB	CSLBP	−0.5%	4.0%	−1.5%	0	**50.0%**
	ORB-CSLBP	12.6%	12.2%	16.7%	50%	**75.0%**
FREAK	CSLBP	5.4%	10.7%	12.2%	18.3%	**19.0%**
	FREAK-CSLBP	5.8%	11.0%	17.8%	36.8%	**63.5%**
KAZE	CSLBP	7.1%	9.7%	**24.5%**	−64.1%	−43.4%
	KAZE-CSLBP	5.8%	13.0%	20.1%	7.6%	**37.9%**
AKAZE	CSLBP	0.1%	0.1%	**1.3%**	−4.0%	−22.4%
	AKAZE-CSLBP	0.1%	1.8%	4.7%	10.4%	**16.8%**

3.5. Influences of Blur and Illumination Changes

With respect to the Bikes sequence (Figure 5), the matching results are shown in Figure 11, Figure 12, Tables 3 and 4. From those figures and tables, we can get the following conclusions:

1. In most scenes, the recall of the CSLBP descriptors is approaching zero and far lower than that of the original descriptors, since the CSLBP feature is sensitive to blur changes. While the combined descriptors are just a small lower.
2. The precision of the CSLBP descriptors is approaching zero and far lower than that of the original descriptors, while combined descriptors are much higher.
3. Among the combined descriptors, the SURF-CSLBP descriptor demonstrates the best matching performance with a less reduced recall (within 9.0%) and a higher precision (over 90%).

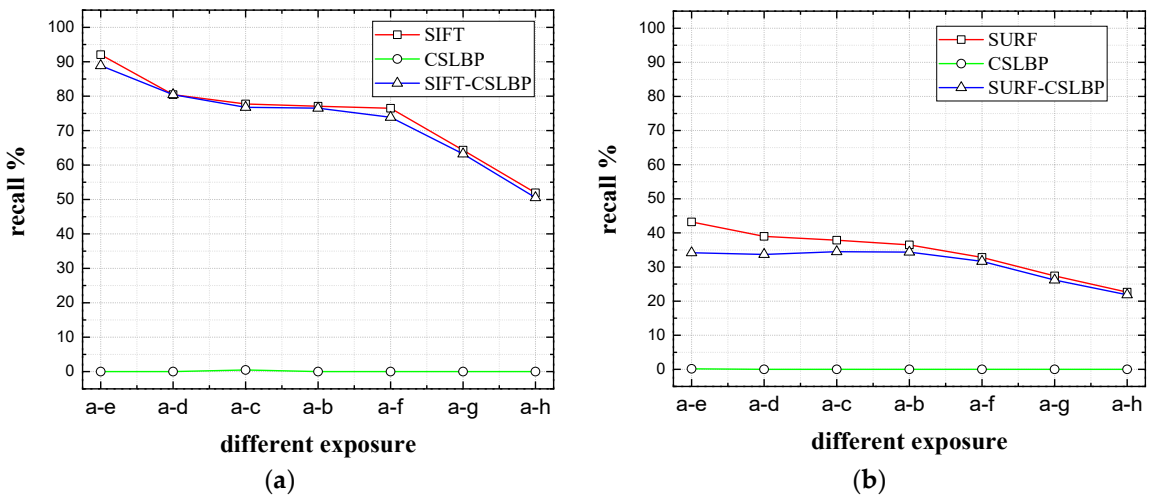


Figure 11. *Cont.*

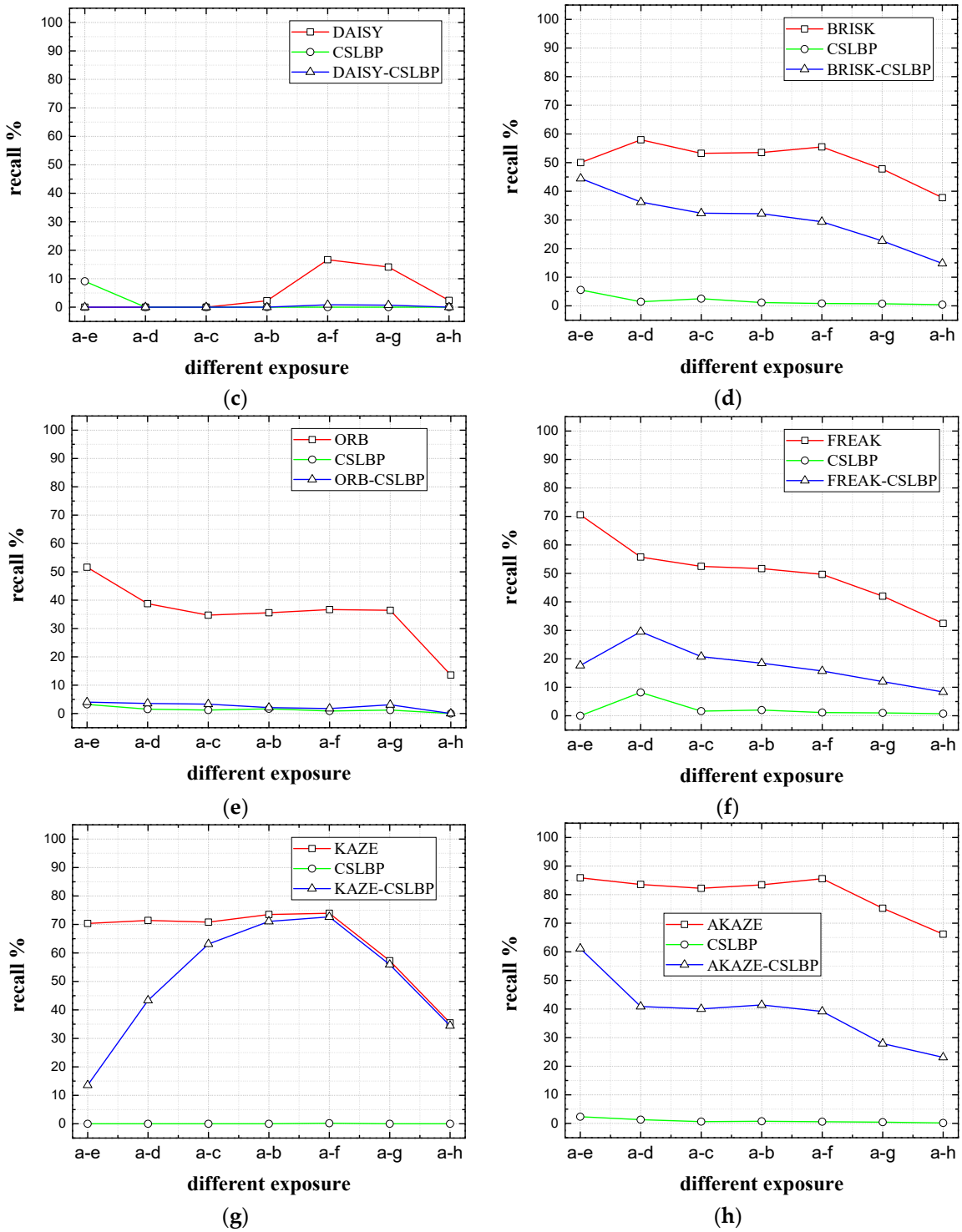


Figure 11. Matching results of the recall of Bikes sequence for the eight original, the CSLBP, and the improved descriptors. The X label represents the matched images of Bikes sequence, and the Y label denotes the recall for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

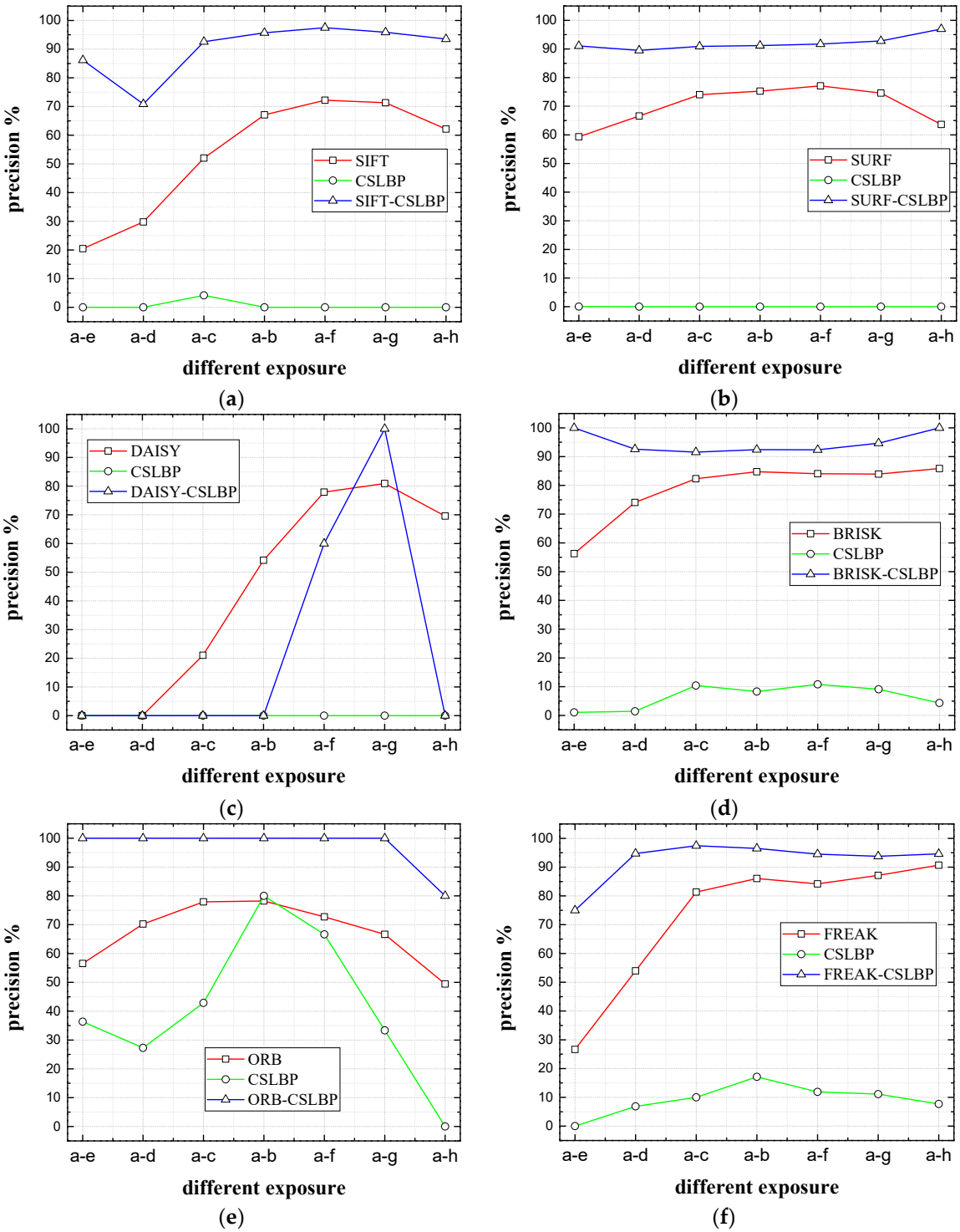
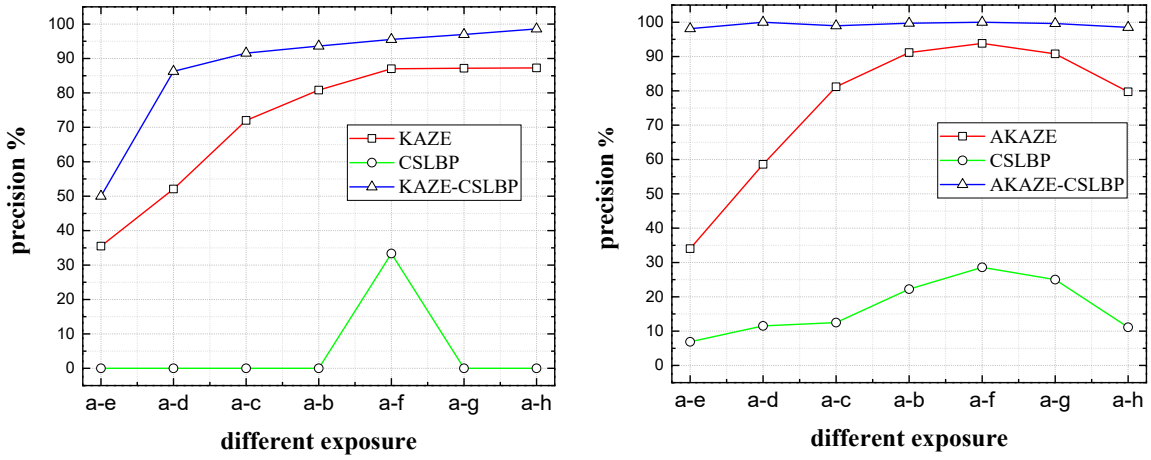


Figure 12. *Cont.*



(**g**) (**h**)

Figure 12. Matching results of the precision of Bikes sequence for the eight original, the CSLBP, and the improved descriptors. The X label represents the matched images of Bikes sequence, and the Y label denotes the precision for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

Table 3. The reduced values of recall for the CSLBP and the combined descriptors compared to the original descriptors of Bikes sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	**−92.1%**	−80.5%	−77.3%	−77.1%	−76.5%	−64.3%	−51.9%
	SIFT-CSLBP	**−3.2%**	0	−0.9%	−0.6%	−2.6%	−1.1%	−1.4%
SURF	CSLBP	**−43.1%**	−39.0%	−37.9%	−36.5%	−32.8%	−27.4%	−22.7%
	SURF-CSLBP	**−9.0%**	−5.3%	−3.4%	−2.1%	−1.1%	−1.2%	−0.8%
DAISY	CSLBP	9.1%	0	0	−2.2%	**−16.7%**	−14.1%	−2.4%
	DAISY-CSLBP	0	0	0	−2.2%	**−15.8%**	−13.4%	−2.4%
BRISK	CSLBP	−44.4%	**−56.5%**	−50.7%	−52.3%	−54.7%	−47.1%	−37.3%
	BRISK-CSLBP	−5.6%	−21.7%	−20.9%	−21.3%	**−26.1%**	−25.1%	−23.0%
ORB	CSLBP	**−48.4%**	−37.2%	−33.5%	−33.9%	−35.8%	−35.2%	−13.6%
	ORB-CSLBP	**−47.6%**	−35.2%	−31.4%	−33.5%	−34.9%	−33.3%	−13.6%
FREAK	CSLBP	**−70.6%**	−47.5%	−50.8%	−49.7%	−48.5%	−41.0%	−31.7%
	FREAK-CSLBP	**−52.9%**	−26.2%	−31.7%	−33.2%	−33.9%	−30.0%	−24.1%
KAZE	CSLBP	−70.4%	−71.4%	−70.8%	−73.5%	**−73.8%**	−57.3%	−35.4%
	KAZE-CSLBP	**−56.8%**	−28.1%	−7.7%	−2.4%	−1.3%	−1.3%	−1.0%
AKAZE	CSLBP	−83.5%	−82.2%	−81.6%	−82.7%	**−85.0%**	−74.8%	−66.0%
	AKAZE-CSLBP	−24.7%	−42.7%	−42.2%	−42.0%	−46.4%	**−47.3%**	−43.1%

Table 4. The improved values of precision for the CSLBP and the combined descriptors compared to the original descriptors of Bikes sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	**−20.4%**	−29.8%	−47.9%	−67.1%	−72.2%	−71.3%	−62.2%
	SIFT-CSLBP	**65.7%**	41.1%	40.5%	28.6%	25.3%	24.5%	31.4%
SURF	CSLBP	**−59.3%**	−66.6%	−74.0%	−75.3%	−77.1%	−74.6%	−63.6%
	SURF-CSLBP	31.7%	22.9%	16.9%	15.9%	14.6%	18.2%	**33.3%**
DAISY	CSLBP	**0.1%**	0	−21.1%	−54.2%	−77.9%	−80.9%	−69.6%
	DAISY-CSLBP	**0**	0	−21.1%	−54.2%	−17.9%	19.1%	−69.6%
BRISK	CSLBP	**−55.2%**	−72.6%	−71.9%	−76.4%	−73.2%	−74.8%	−81.5%
	BRISK-CSLBP	**43.8%**	18.5%	9.2%	7.7%	8.3%	10.7%	14.1%
ORB	CSLBP	−20.2%	−43.0%	−35.1%	**1.8%**	−6.1%	−33.3%	−49.5%
	ORB-CSLBP	**43.5%**	29.8%	22.1%	21.8%	27.3%	33.3%	30.5%
FREAK	CSLBP	**−26.7%**	−47.1%	−71.4%	−68.9%	−72.3%	−76.0%	−83.0%
	FREAK-CSLBP	**48.3%**	40.8%	16.1%	10.5%	10.4%	6.6%	3.9%
KAZE	CSLBP	**−35.5%**	−52.1%	−72.0%	−80.8%	−53.7%	−87.2%	−87.2%
	KAZE-CSLBP	**14.5%**	34.2%	19.5%	12.8%	8.5%	9.8%	11.3%
AKAZE	CSLBP	**−27.1%**	−47.1%	−68.7%	−68.9%	−65.3%	−65.8%	−68.6%
	AKAZE-CSLBP	**64.1%**	41.4%	17.7%	8.5%	6.2%	8.8%	18.8%

3.6. Influences of Scale, Rotation and Illumination Changes

As for Boat sequence (Figure 6), the matching results are shown in Figure 13, Figure 14, Tables 5 and 6. From those figures and tables, we can get the following conclusions:

1. In most scenes, the recall of the CSLBP descriptors is approaching zero and far lower than that of the original descriptors except that the recall of DAISY is zero as well, since the CSLBP feature is sensitive to scale changes. While the combined descriptors are mostly higher.
2. The precision of the CSLBP descriptors mostly approaching zero except that of the CSLBP descriptor for ORB features and far lower than that of the original descriptors except that the precision of DAISY is zero as well. While the combined descriptors are mostly much higher except that the precision of DAISY-CSLBP is still zero.
3. Among the combined descriptors, the ORB-CSLBP descriptor shows the best matching performance with a less reduced recall (within 9.7%) and a higher precision (over 90%).

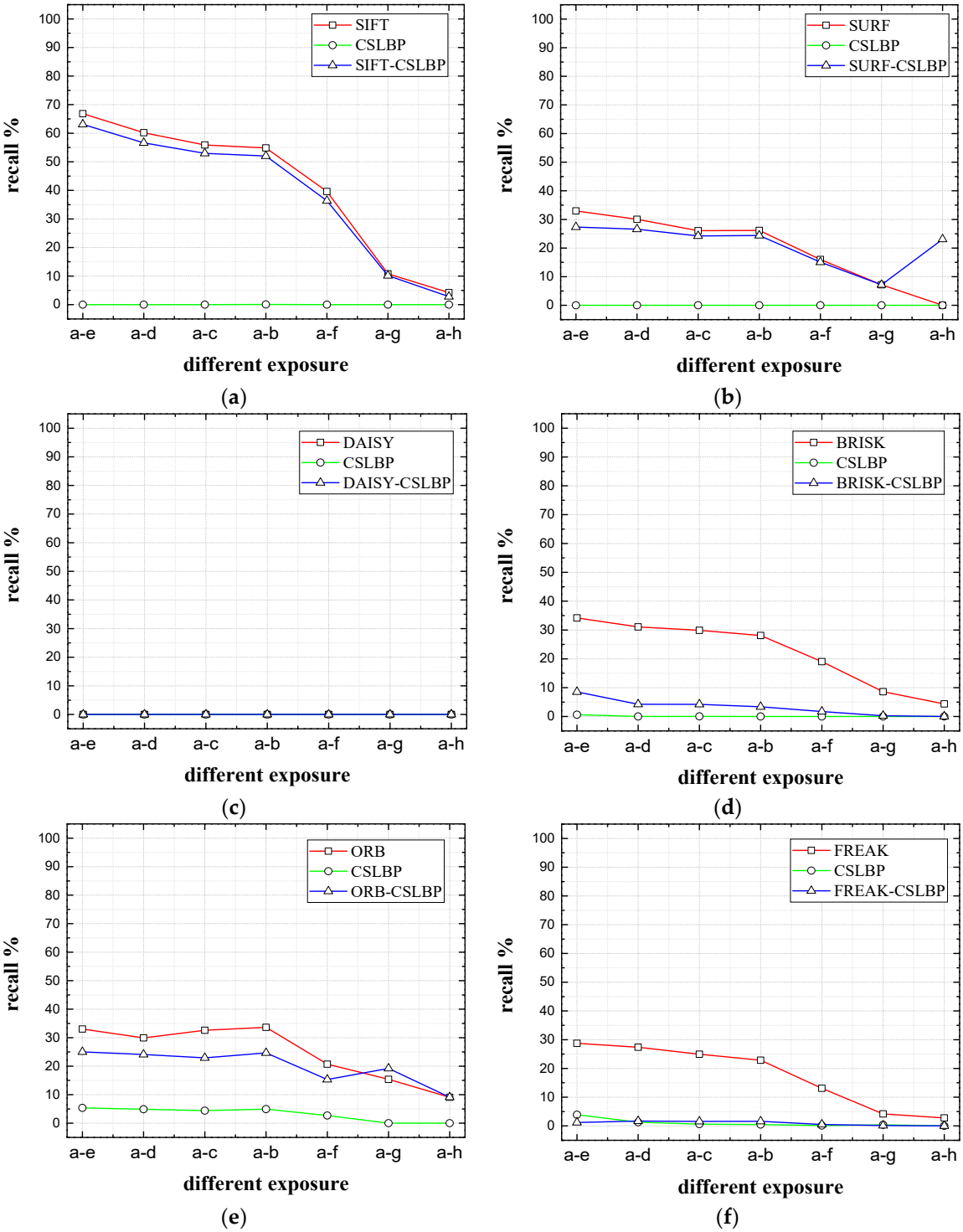


Figure 13. *Cont.*

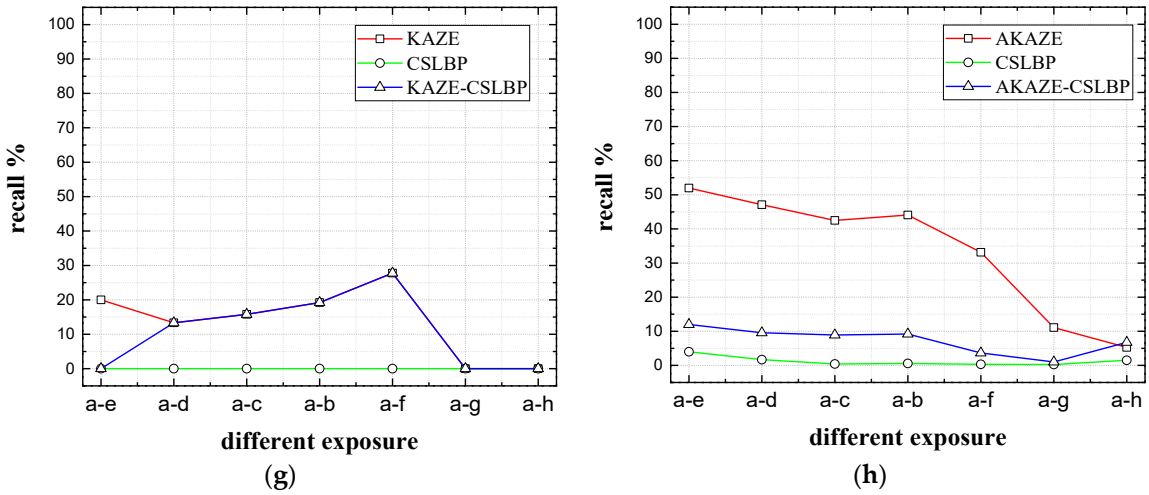


Figure 13. Matching results of the recall of Boat sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Boat sequence, and the Y label denotes the recall for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

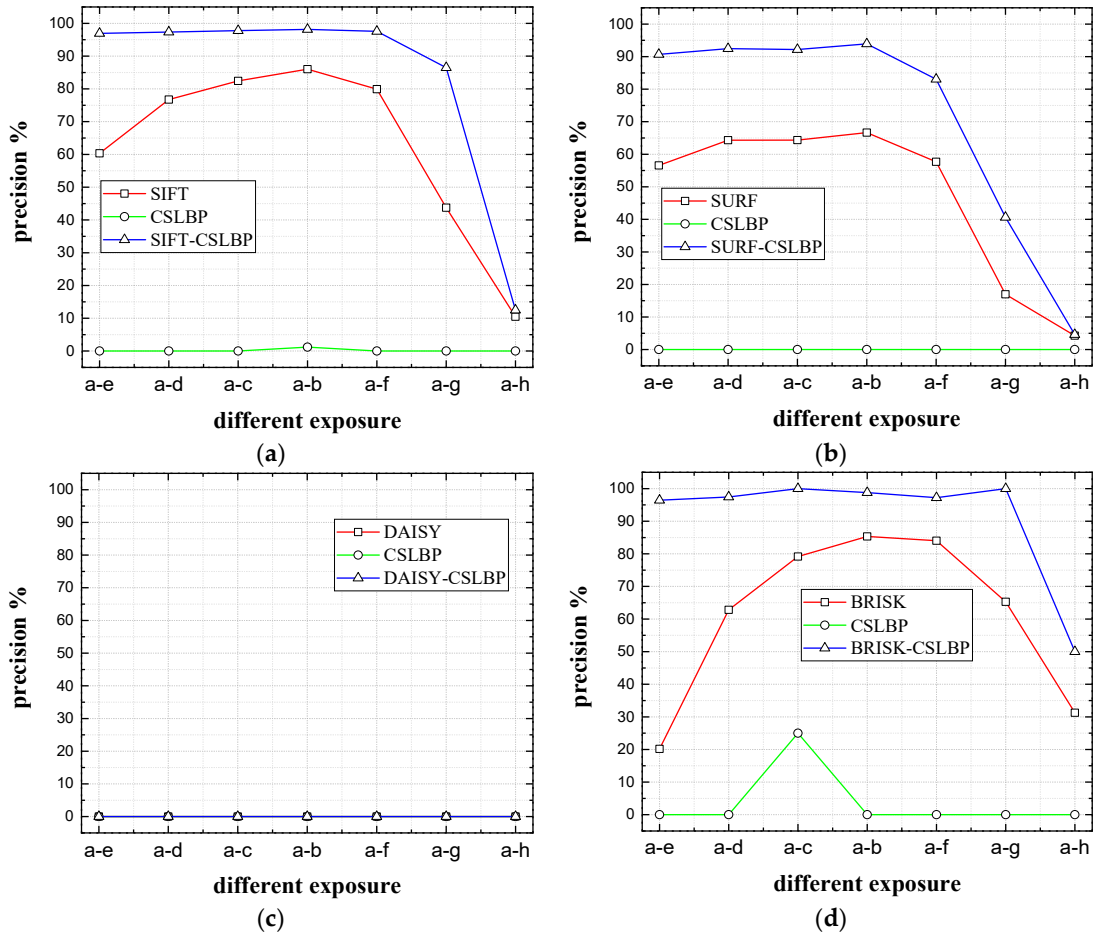


Figure 14. *Cont.*

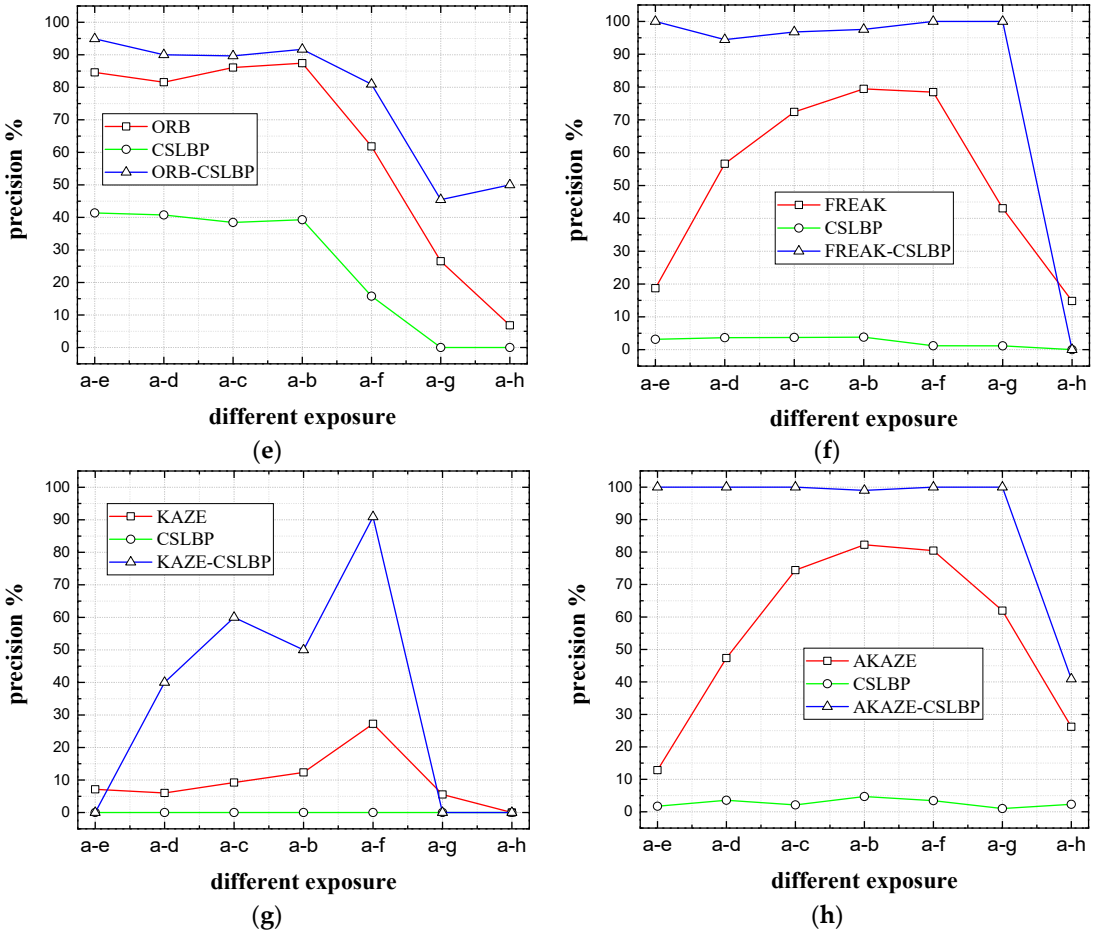


Figure 14. Matching results of the precision of Boat sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Boat sequence, and the Y label denotes the precision for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; and (**h**) AKAZE features.

Table 5. The reduced values of recall for the CSLBP and the combined descriptors compared to the original descriptors of Boat sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	**−66.8%**	−60.2%	−55.9%	−54.8%	−39.6%	−10.8%	−4.2%
	SIFT-CSLBP	**−3.7%**	−3.6%	−2.9%	−2.8%	−3.2%	−0.6%	−1.4%
SURF	CSLBP	**−33.0%**	−30.1%	−26.1%	−26.2%	−16.0%	−7.2%	0
	SURF-CSLBP	**−5.6%**	−3.5%	−1.8%	−1.8%	−0.9%	0	23.1%
DAISY	CSLBP	0	0	0	0	0	0	0
	DAISY-CSLBP	0	0	0	0	0	0	0
BRISK	CSLBP	**−33.5%**	−31.1%	−29.8%	−28.1%	−19.0%	−8.6%	−4.4%
	BRISK-CSLBP	−25.6%	**−26.8%**	−25.7%	−24.7%	−17.3%	−8.4%	−4.4%
ORB	CSLBP	**−27.7%**	−25.0%	−28.2%	−28.7%	−18.0%	−15.4%	−9.1%
	ORB-CSLBP	−8.0%	−5.8%	**−9.7%**	−9.0%	−5.4%	3.8%	0
FREAK	CSLBP	**−24.9%**	−26.1%	−24.3%	−22.4%	−12.9%	−3.8%	−2.8%
	FREAK-CSLBP	**−27.5%**	−25.7%	−23.4%	−21.3%	−12.6%	−4.0%	−2.8%
KAZE	CSLBP	−20.0%	−13.3%	−15.8%	−19.2%	**−27.8%**	0	0
	KAZE-CSLBP	**−20.0%**	0	0	0	0	0	0
AKAZE	CSLBP	**−48.0%**	−45.4%	−42.1%	−43.5%	−32.8%	−10.8%	−3.8%
	AKAZE-CSLBP	**−40.0%**	−37.5%	−33.6%	−34.9%	−29.5%	−10.1%	1.5%

Table 6. The improved values of precision for the CSLBP and the combined descriptors compared to the original descriptors of Boat sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	−60.3%	−76.7%	−82.4%	−84.8%	−79.9%	−43.8%	**−10.5%**
	SIFT-CSLBP	**36.6%**	20.6%	15.3%	12.1%	17.6%	42.7%	2.0%
SURF	CSLBP	−56.6%	−64.3%	−64.3%	−66.7%	−57.7%	−17.0%	**−4.3%**
	SURF-CSLBP	**34.1%**	28.1%	27.8%	27.3%	25.4%	23.6%	0.3%
DAISY	CSLBP	0	0	0	0	0	0	0
	DAISY-CSLBP	**0**	0	0	0	0	0	**0**
BRISK	CSLBP	**−20.2%**	−62.8%	−54.2%	−85.3%	−84.0%	−65.3%	−31.3%
	BRISK-CSLBP	**76.2%**	34.6%	20.8%	13.4%	13.2%	34.7%	18.8%
ORB	CSLBP	−43.2%	−40.8%	−47.6%	−48.1%	−46.0%	−26.5%	**−6.8%**
	ORB-CSLBP	10.3%	8.5%	3.6%	4.2%	19.1%	18.9%	**43.2%**
FREAK	CSLBP	−15.5%	−53.0%	−68.7%	−75.6%	−77.3%	−41.9%	**−14.8%**
	FREAK-CSLBP	**81.3%**	37.8%	24.4%	18.1%	21.5%	57.0%	35.2%
KAZE	CSLBP	−7.1%	−6.0%	−9.2%	−12.3%	−27.3%	−5.6%	**0**
	KAZE-CSLBP	−7.1%	34.0%	50.8%	37.7%	**63.6%**	−5.6%	0
AKAZE	CSLBP	**−11.1%**	−43.8%	−72.3%	−77.5%	−77.0%	−60.9%	−23.9%
	AKAZE-CSLBP	**87.2%**	52.6%	25.6%	16.7%	19.6%	38.1%	14.7%

3.7. Influences of Viewpoint and Illumination Changes

For the Graffiti sequence (Figure 7), the matching results are shown in Figure 15, Figure 16, Tables 7 and 8. From those figures and tables, we can get the following conclusions:

1. In most scenes, the recall of the CSLBP descriptors is approaching zero and far lower than that of the original descriptors except that the recall of DAISY and KAZE is approaching zero as well, since the CSLBP feature is sensitive to viewpoint changes. While the combined descriptors are mostly higher.
2. The precision of the CSLBP descriptors is mostly approaching zero except that of the CSLBP descriptor for BRISK and FREAK features and far lower than that of the original descriptors except that the precision of DAISY and KAZE descriptors is almost zero as well. While the combined descriptors are mostly obviously higher except that the precision of DAISY-CSLBP and KAZE-CSLBP descriptors is still nearly zero.
3. Among the combined descriptors, the FREAK-CSLBP descriptor shows the best matching performance with a less reduced recall (within 12.6%) and a higher precision (over 50%).

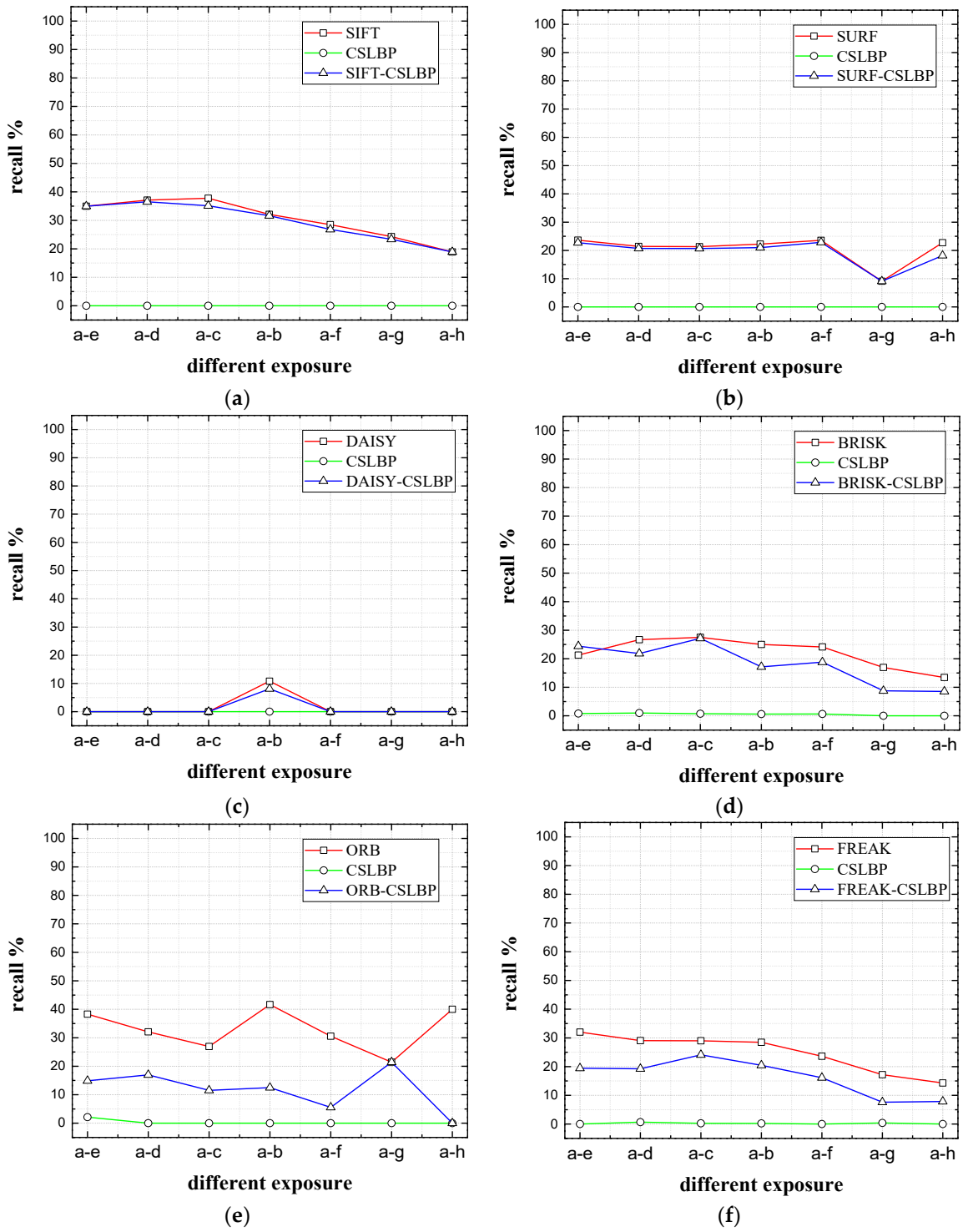
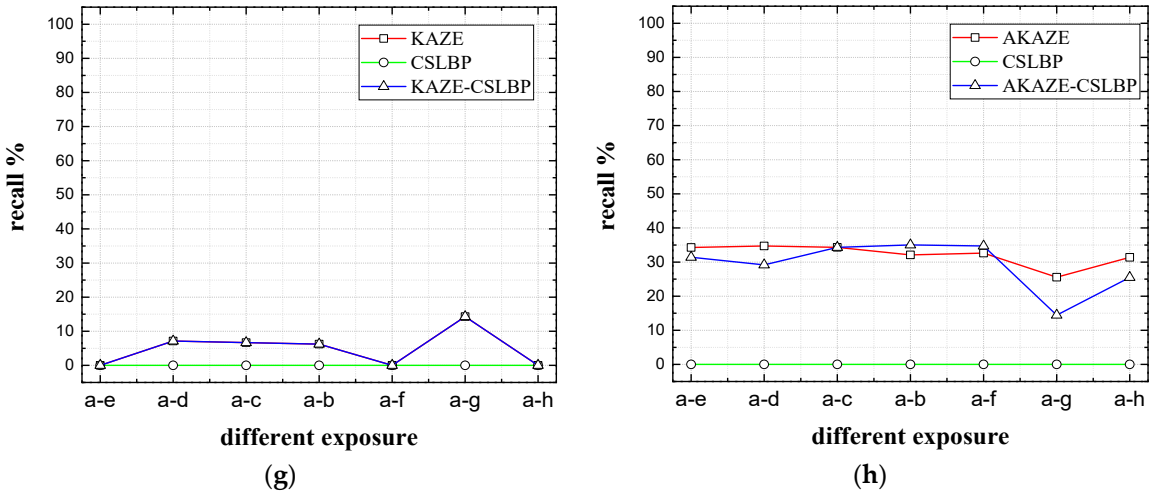
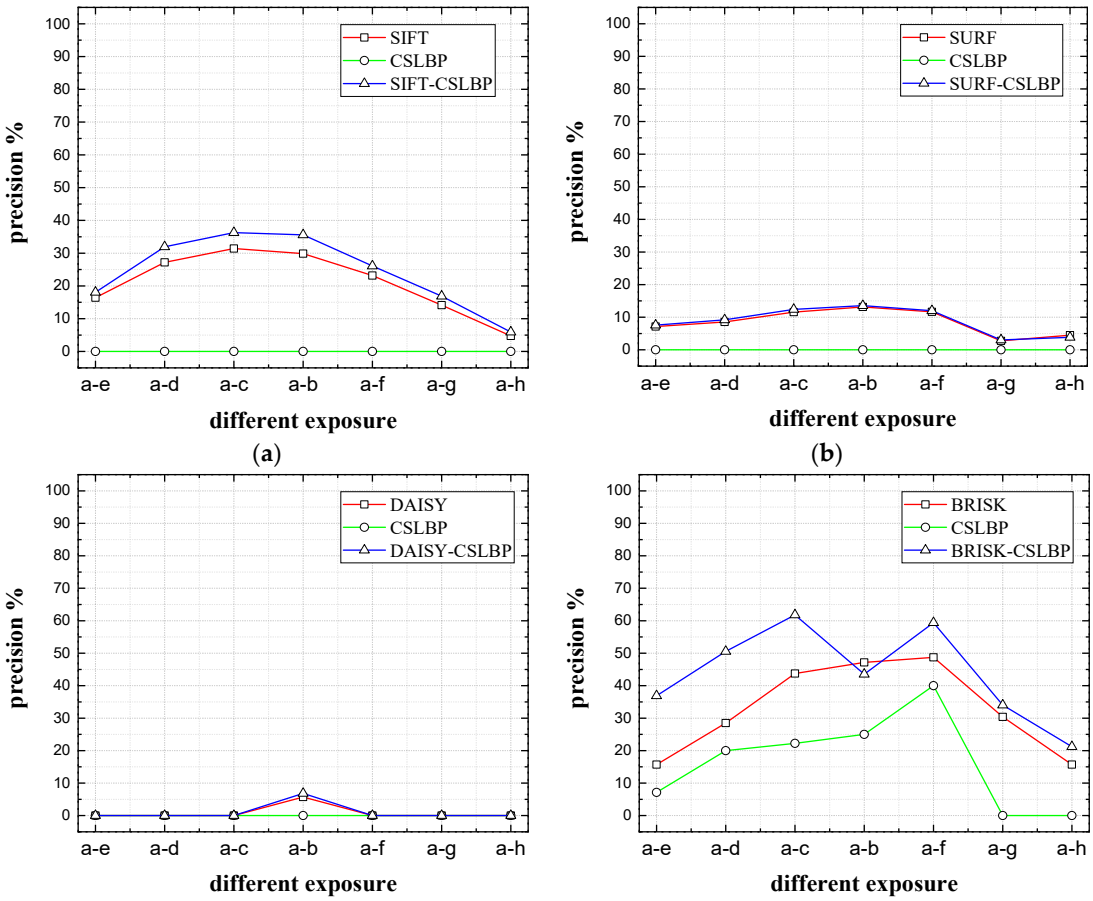


Figure 15. *Cont.*



(g) (h)

Figure 15. Matching results of the recall of Graffiti sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Graffiti sequence, and the Y label denotes the recall for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; and (**h**) AKAZE features.



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

Figure 16. *Cont.*

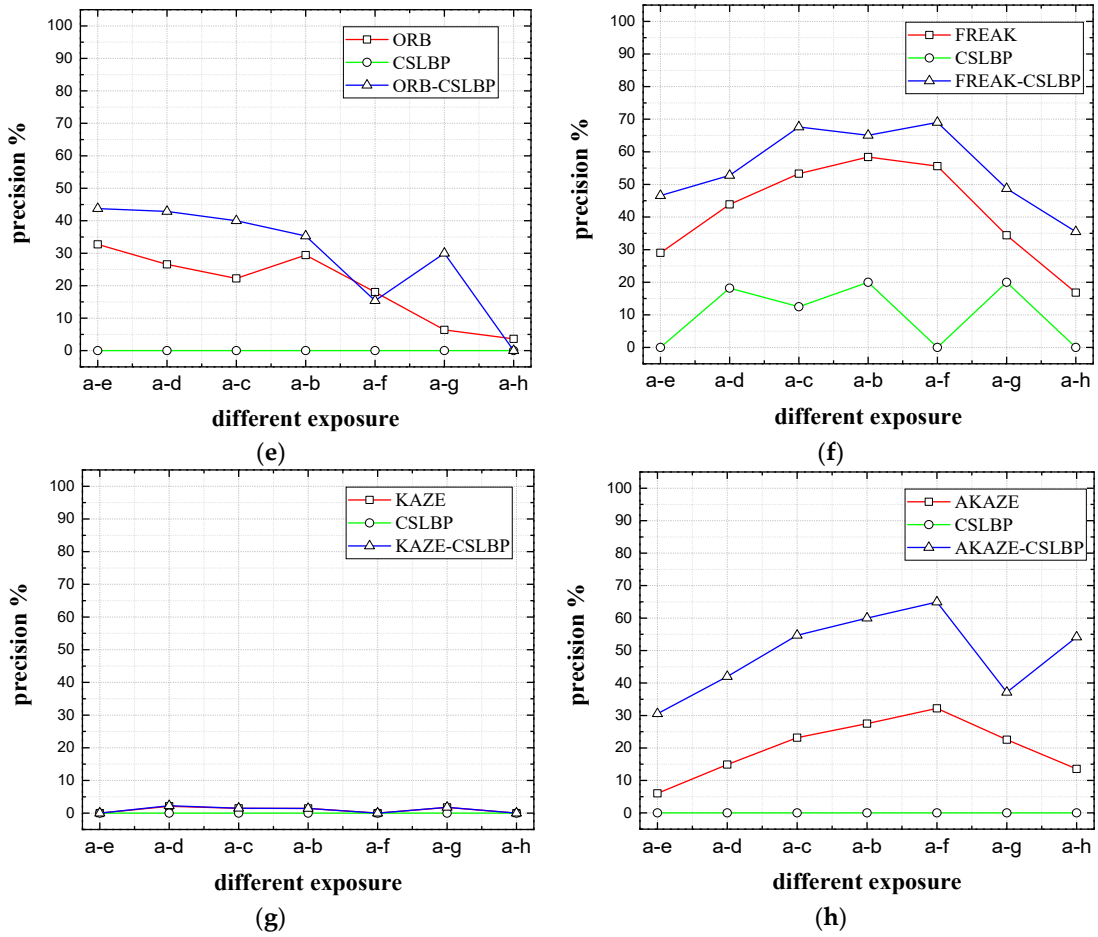


Figure 16. Matching results of the precision of Graffiti sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Graffiti sequence, and the Y label denotes the precision for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

Table 7. The reduced values of recall for the CSLBP and the combined descriptors compared to the original descriptors of Graffiti sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	−35.0%	−37.1%	**−37.8%**	−32.1%	−28.5%	−24.3%	−18.9%
	SIFT-CSLBP	0	−0.6%	**−2.7%**	−0.5%	−1.7%	−0.9%	0
SURF	CSLBP	**−23.6%**	−21.4%	−21.3%	−22.2%	−23.6%	−9.1%	−22.7%
	SURF-CSLBP	−0.9%	−0.7%	−0.7%	−1.2%	−0.7%	0	**−4.5%**
DAISY	CSLBP	0	0	0	**−10.8%**	0	0	0
	DAISY-CSLBP	0	0	0	**−2.7%**	0	0	0
BRISK	CSLBP	−20.5%	−25.7%	**−26.8%**	−24.4%	−23.5%	−16.9%	−13.4%
	BRISK-CSLBP	3.1%	−4.9%	−0.4%	−7.8%	−5.3%	**−8.2%**	−4.9%
ORB	CSLBP	−36.2%	−32.1%	−26.9%	**−41.7%**	−30.6%	−21.4%	−40.0%
	ORB-CSLBP	−23.4%	−15.1%	−15.4%	−29.2%	−25.0%	0	**−40.0%**
FREAK	CSLBP	**−32.0%**	−28.4%	−28.7%	−28.2%	−23.6%	−16.8%	−14.3%
	FREAK-CSLBP	**−12.6%**	−9.8%	−4.9%	−8.0%	−7.5%	−9.6%	−6.4
KAZE	CSLBP	0	−7.1%	−6.7%	−6.3%	0	**−14.3%**	0
	KAZE-CSLBP	0	0	0	0	0	0	0
AKAZE	CSLBP	−34.3%	**−34.7%**	−34.3%	−32.1%	−32.6%	−25.6%	−31.4%
	AKAZE-CSLBP	−2.9%	−5.6%	0	2.9%	2.1%	**−11.1%**	−5.9%

Table 8. The improved values of precision for the CSLBP and the combined descriptors compared to the original descriptors of Graffiti sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	−16.4%	−27.2%	−31.4%	−29.9%	−23.2%	−14.1%	**−4.7%**
	SIFT-CSLBP	1.7%	4.7%	4.8%	**5.7%**	2.9%	2.8%	1.2%
SURF	CSLBP	−7.1%	−8.5%	−11.6%	−13.1%	−11.7%	**−2.8%**	−4.5%
	SURF-CSLBP	0.4%	0.7%	**0.8%**	0.4%	0.3%	0.3%	−0.7%
DAISY	CSLBP	**0**	0	0	−5.1%	0	0	0
	DAISY-CSLBP	0	0	0	**1.1%**	0	0	0
BRISK	CSLBP	−8.6%	**−8.5%**	−21.5%	−22.2%	−8.7%	−30.4%	−15.7%
	BRISK-CSLBP	21.2%	**22.1%**	18.0%	−3.6%	10.7%	3.7%	5.5%
ORB	CSLBP	−32.7%	−26.6%	−22.2%	−29.4%	−18.0%	−6.4%	**−3.6%**
	ORB-CSLBP	11.0%	16.3%	17.8%	5.9%	−2.6%	**23.6%**	−3.6%
FREAK	CSLBP	−29.0%	−25.7%	−40.8%	−38.4%	−55.6%	**−14.4%**	−16.8%
	FREAK-CSLBP	17.6%	8.9%	14.3%	6.7%	13.4%	14.3%	**18.7%**
KAZE	CSLBP	**0**	−2.1%	−1.5%	−1.4%	0	−1.8%	0
	KAZE-CSLBP	0	**0.2%**	0.1%	0.1%	0	0.1%	0
AKAZE	CSLBP	**−6.0%**	−14.9%	−23.2%	−27.5%	−32.2%	−22.5%	−13.6%
	AKAZE-CSLBP	24.5%	27.1%	31.5%	32.5%	32.7%	14.6%	**40.6%**

3.8. Influences of JPEG Compression and Illumination Changes

With respect to Ubc sequence (Figure 8), the matching results are shown in Figure 17, Figure 18, Tables 9 and 10. From those figures and tables, we can get the following conclusions:

1. In most scenes, the recall of the CSLBP descriptors is far lower than that of the original descriptors and that of the non-binary CSLBP descriptors for SIFT, SURF, DAISY and KAZE features is almost approaching zero while the binary CSLBP descriptors for BRISK, ORB, FREAK and AKAZE features not, since the binary CSLBP feature is more robustness against the JPEG compression changes than the non-binary CSLBP feature. While the combined descriptors are mostly higher.
2. The precision of the CSLBP descriptor is mostly lower than that of the original descriptors, while the combined descriptor is all much higher.
3. Among the combined descriptors, the ORB-CSLBP descriptor shows the best matching performance with a less reduced recall (within 20.8%) and a higher precision (over 70%).

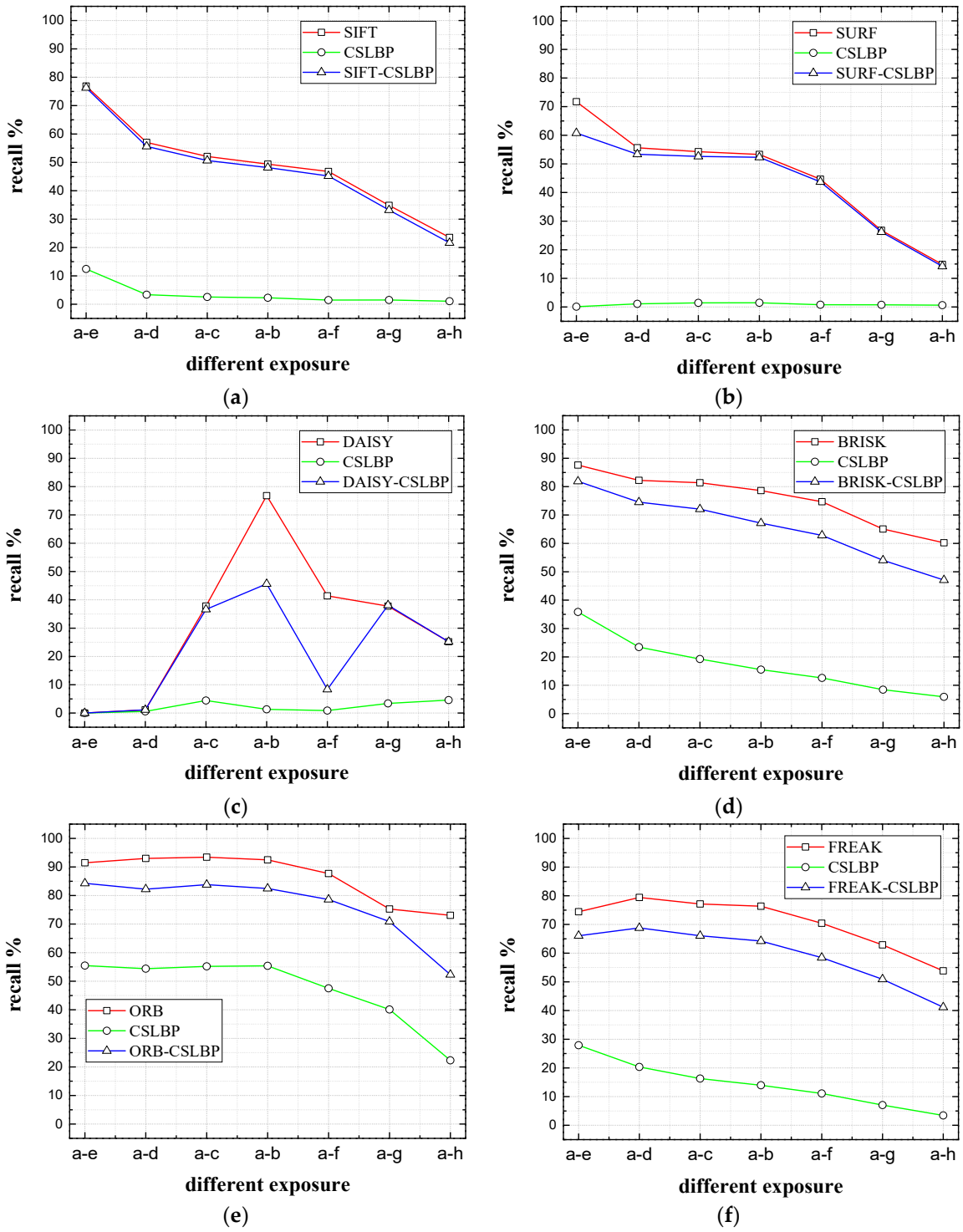


Figure 17. *Cont.*

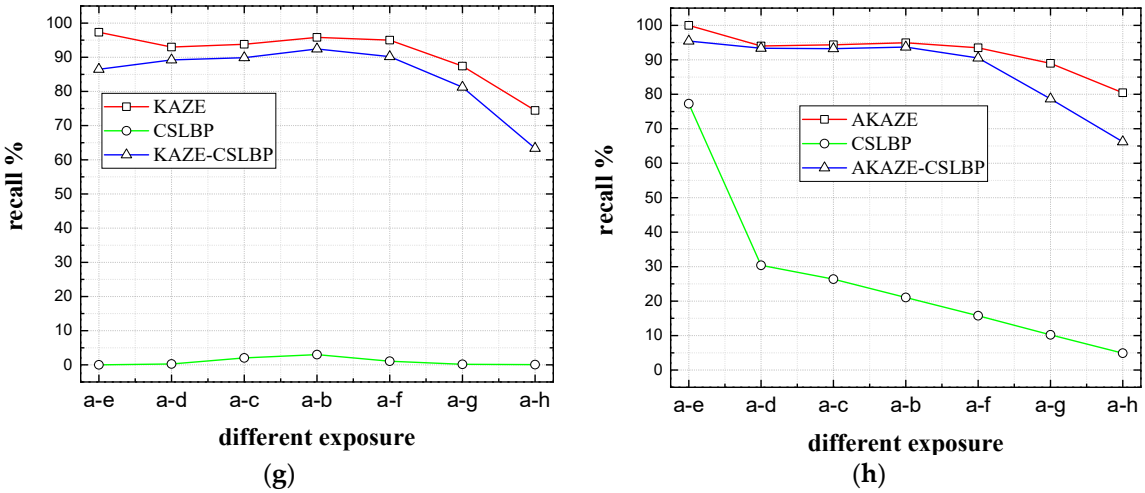


Figure 17. Matching results of the recall of Ubc sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Ubc sequence, and the Y label denotes the recall for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

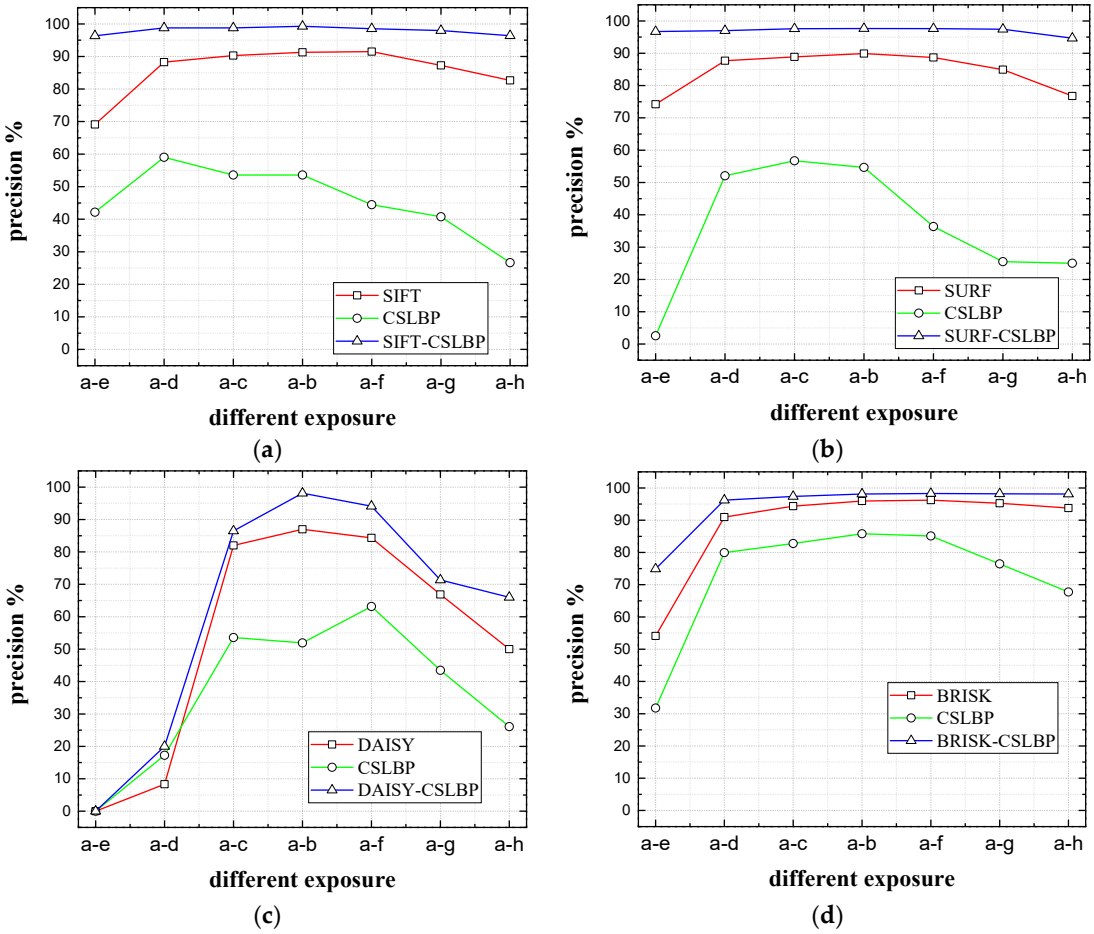


Figure 18. *Cont.*

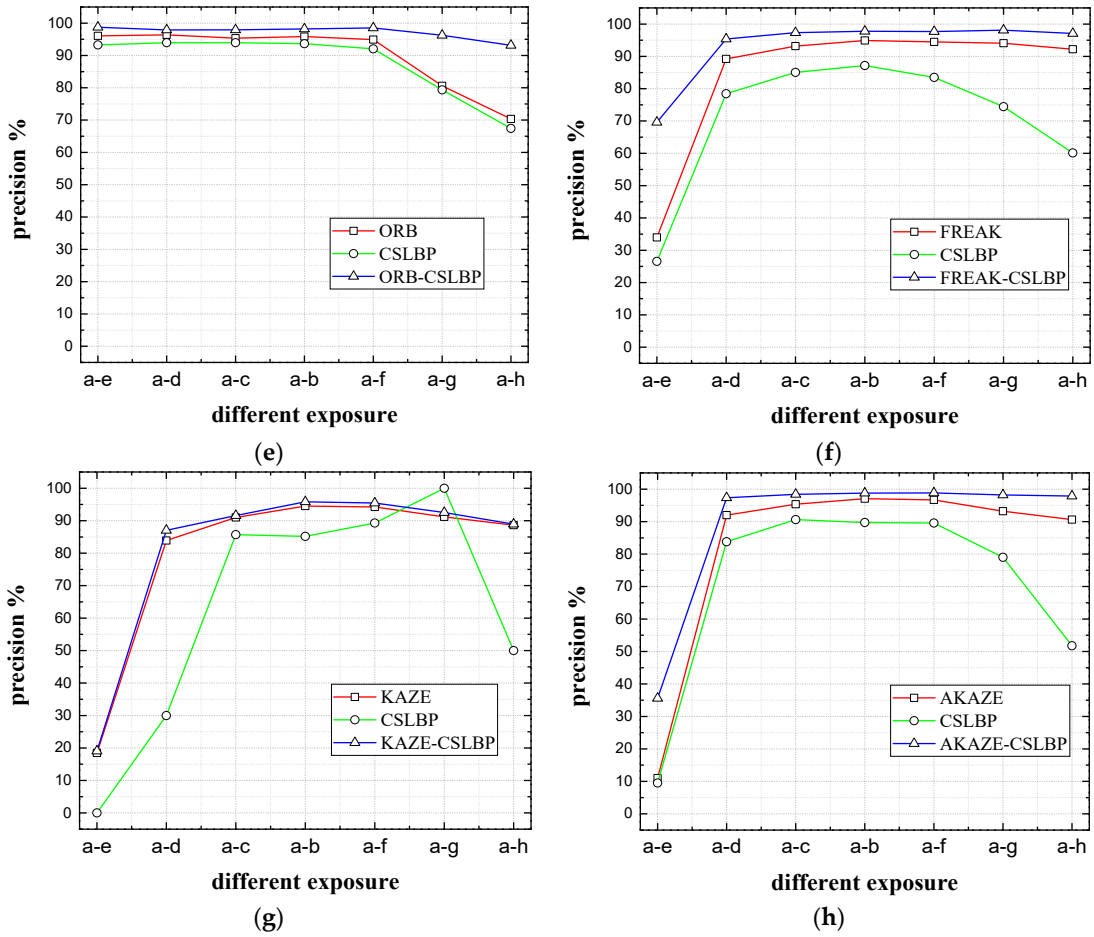


Figure 18. Matching results of the precision of Ubc sequence for the eight original, the CSLBP, and the combined descriptors. The X label represents the matched images of Ubc sequence, and the Y label denotes the precision for each descriptor. (**a**) SIFT features; (**b**) SURF features; (**c**) DAISY features; (**d**) BRISK features; (**e**) ORB features; (**f**) FREAK features; (**g**) KAZE features; (**h**) AKAZE features.

Table 9. The reduced values of recall for the CSLBP and the combined descriptors compared to the original descriptors of Ubc sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	**−64.5%**	−53.7%	−49.5%	−47.1%	−45.3%	−33.3%	−22.5%
	SIFT-CSLBP	−0.6%	−1.4%	−1.4%	−1.2%	−1.6%	−1.6%	**−1.9%**
SURF	CSLBP	**−71.6%**	−54.5%	−52.8%	−51.8%	−43.9%	−26.0%	−14.2%
	SURF-CSLBP	**−10.9%**	−2.3%	−1.6%	−1.0%	−0.9%	−0.6%	−0.6%
DAISY	CSLBP	0	−0.6%	−33.4	**−75.5%**	−40.5%	−34.4%	−20.6%
	DAISY-CSLBP	0	0	−1.2%	−31.1%	**−33.1%**	0.3%	0
BRISK	CSLBP	−51.8%	−58.8%	−62.1%	**−63.1%**	−62.1%	−56.6%	−54.3%
	BRISK-CSLBP	−5.8%	−7.7%	−9.3%	−11.5%	−11.9%	−11.0%	**−13.1%**
ORB	CSLBP	−36.0%	−38.6%	−38.2%	−37.1%	−40.2%	−35.2%	**−50.8%**
	ORB-CSLBP	−7.2%	−10.8%	−9.6%	−10.0%	−9.1%	−4.4	**−20.8%**
FREAK	CSLBP	−46.5%	−59.1%	−60.9%	**−62.4%**	−59.3%	−55.8%	−50.4%
	FREAK-CSLBP	−8.4%	−10.6%	−11.1%	−12.1%	−12.0%	−12.0%	**−12.7%**
KAZE	CSLBP	**−97.3%**	−92.7%	−91.8%	−92.8%	−93.9%	−87.3%	−74.4%
	KAZE-CSLBP	−10.8%	−3.8%	−3.9%	−3.4%	−4.8%	−6.2%	**−11.1%**
AKAZE	CSLBP	−22.7%	−63.6%	−68.0%	−73.9%	−77.7%	**−78.7%**	−75.6%
	AKAZE-CSLBP	−4.5%	−0.6%	−1.1%	−1.2%	−3.0%	−10.3%	**−14.2%**

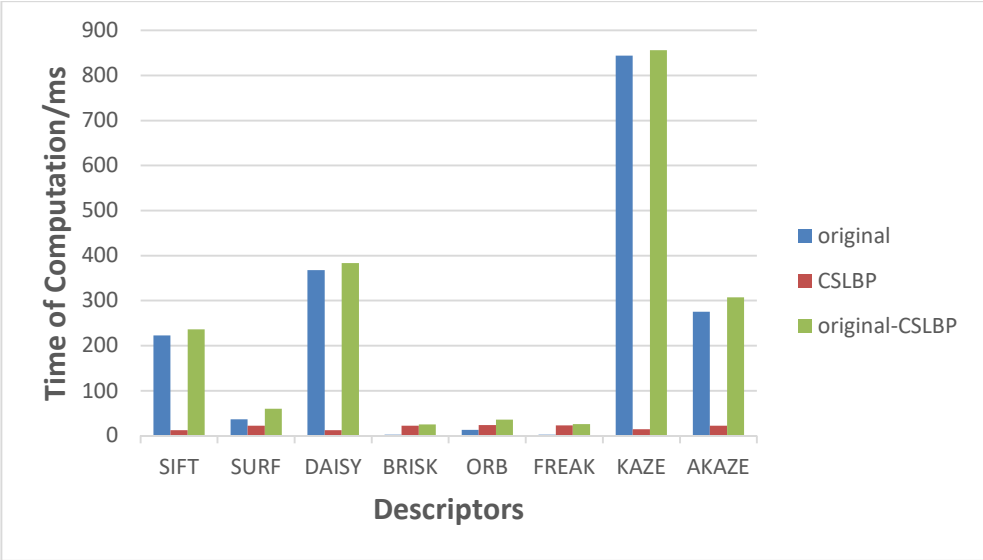
Table 10. The improved values of precision for the CSLBP and the combined descriptors compared to the original descriptors of Ubc sequence.

Features	Descriptors	a-e	a-d	a-c	a-b	a-f	a-g	a-h
SIFT	CSLBP	**−27.0%**	−29.2%	−36.7%	−37.7%	−47.0%	−46.5%	−56.0%
	SIFT-CSLBP	**27.2%**	10.5%	8.5%	8.0%	7.0%	10.7%	13.7%
SURF	CSLBP	−71.6%	−35.6%	**−32.2%**	−35.2%	−52.3%	−59.4%	−51.8%
	SURF-CSLBP	**22.5%**	9.3%	8.7%	7.7%	8.9%	12.5%	17.9%
DAISY	CSLBP	0	**8.9%**	−28.4%	−35.1%	−21.2%	−23.4%	−23.9%
	DAISY-CSLBP	0	11.7%	4.5%	11.1%	9.8%	4.5%	**16.0%**
BRISK	CSLBP	−22.3%	−11.0%	−11.6%	**−10.1%**	−11.1%	−18.8%	−26.0%
	BRISK-CSLBP	**20.8%**	5.3%	3.0%	2.2%	2.1%	2.9%	4.4%
ORB	CSLBP	−2.8%	−2.4%	−1.4%	−2.2%	−2.9%	**−1.2%**	−2.9%
	ORB-CSLBP	2.7%	1.5%	2.6%	2.4%	3.6%	15.7%	**22.8%**
FREAK	CSLBP	**−7.4%**	−10.8%	−8.1%	−7.7%	−11.0%	−19.6%	−32.1%
	FREAK-CSLBP	**35.6%**	6.2%	4.2%	2.9%	3.2%	4.1%	4.9%
KAZE	CSLBP	−18.5%	−53.9%	−5.3%	−9.3%	−4.9%	**8.8%**	−38.6%
	KAZE-CSLBP	0.6%	**3.2%**	0.6%	1.3%	1.2%	1.3%	0.3%
AKAZE	CSLBP	**−1.5%**	−8.2%	−4.8%	−7.3%	−7.1%	−14.2%	−38.8%
	AKAZE-CSLBP	**24.6%**	5.3%	3.0%	1.7%	2.1%	5.0%	7.3%

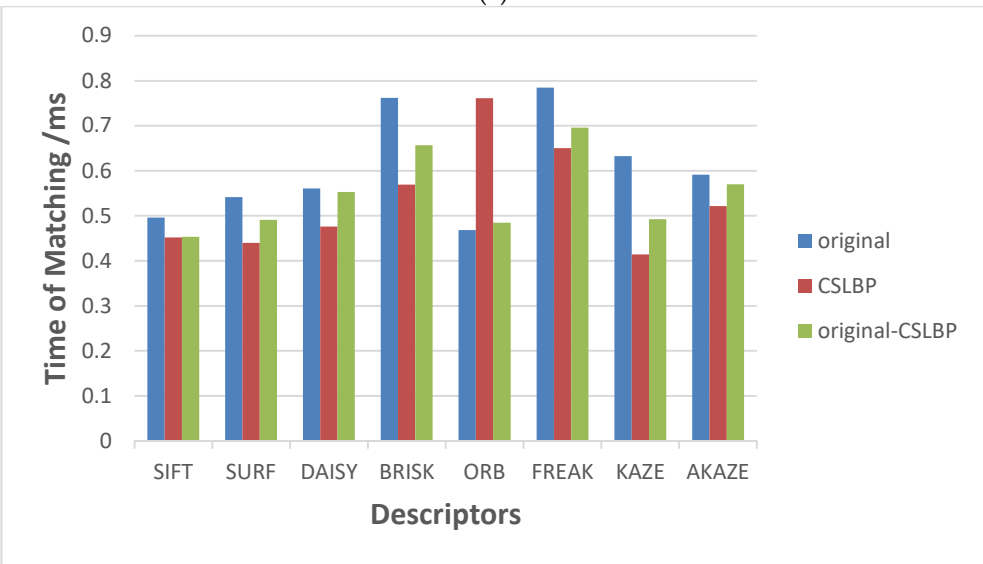
3.9. Computational and Matching Speed Analysis

The time consumption of construction and matching for each descriptor of 30 keypoints are shown in Figure 19a,b. The ratios of the CSLBP and combined descriptors to the original descriptors in computational and matching speed are computed in Table 11. From those figures and table, we can get the following conclusions:

1. The computational speed of CSLBP descriptor is much faster than that of SIFT, DAISY, KAZE and AKAZE descriptor and a bit faster than that of SURF descriptor, while slower than that of BRISK, ORB and FREAK descriptor. While that of SIFT-CSLBP, DAISY-CSLBP, KAZE-CSLBP and AKAZE-CSLBP descriptor is just slightly slower than that of SIFT, DAISY, KAZE and AKAZE descriptor and that of SURF-CSLBP descriptor is a bit lower than that of SURF descriptor, that of BRISK-CSLBP, ORB-CSLBP and FREAK-CSLBP descriptors is much slower than that of BRISK, ORB and FREAK descriptor.
2. The matching speed of CSLBP descriptor is a little faster than that of SIFT, SURF, DAISY, BRISK, FREAK, KAZE, and AKAZE descriptor, while a bit slower than that of ORB descriptor. Further, the improved descriptors is also a bit faster than that of the original descriptors except that the matching speed of ORB-CSLBP descriptor is slightly slower than that of ORB descriptor.
3. Since the computational time consumption for the descriptor is much longer than the matching time consumption, the total time consumption of the improved algorithms is a litter longer than that of the original algorithms.



(**a**)



(**b**)

Figure 19. Comparison results of computational and matching time consumption for the eight original, the CSLBP and the improved descriptors. The blue histograms represent the time consumption of the original descriptors, the red histograms denote the time consumption of the CSLBP descriptors and the green histograms indicate the time consumption of the improved descriptors. (**a**) Computational time consumption of 30 descriptors; and (**b**) matching time consumption of 30 descriptors against 30 descriptors.

Table 11. The ratios of the CSLBP and combined descriptors to the original descriptors in computational and matching speed.

Features	Descriptors	Computation Speed	Matching Speed
SIFT	CSLBP	17.81	1.10
	SIFT-CSLBP	0.94	1.09
SURF	CSLBP	1.66	1.23
	SURF-CSLBP	0.61	1.10
DAISY	CSLBP	29.25	1.18
	DAISY-CSLBP	0.96	1.02
BRISK	CSLBP	0.10	1.34
	BRISK-CSLBP	0.09	1.16
ORB	CSLBP	0.53	0.62
	ORB-CSLBP	0.36	0.97
FREAK	CSLBP	0.11	1.21
	FREAK-CSLBP	0.10	1.13
KAZE	CSLBP	58.92	1.53
	KAZE-CSLBP	0.99	1.28
AKAZE	CSLBP	12.32	1.13
	AKAZE-CSLBP	0.89	1.04

4. Discussion

This paper focuses on addressing the accuracy problem of eight common feature matching algorithms (SIFT, SURF, DAISY, BRISK, ORB, FREAK, KAZE, and AKAZE). The decrease in accuracy under strong illumination changes leads to the poor performance in AR system. We discovered that the CSLBP descriptor has the characteristics of strong illumination invariance, so we create new descriptors by combining the descriptor of original algorithms with the CSLBP descriptor. Experiments are carried out to prove that the robustness against light changes can be improved when using our combined descriptors. We conduct experiments in five different scenes with varied illumination conditions to find out the performance of these descriptors in different environments. Further, experiments on the time consumption of matching and computation are carried out.

Our experimental results have shown that the precision of feature matching by the combined descriptors has improved significantly in most of the street view scenes with strong illumination changes. Specifically, the FREAK-CSLBP descriptor shows the best matching performance under only illumination changes. The SURF-CSLBP descriptor suggests the best matching performance under blur and illumination changes. The ORB-CSLBP descriptor shows the best matching performance under scale, rotation and illumination changes. The FREAK-CSLBP descriptor suggests the best performance under viewpoint and illumination changes. Finally, the ORB-CSLBP descriptor shows the best matching performance under different JPEG compression and illumination changes. In addition, the time consumption of the combined descriptors is just a litter longer than that of the original algorithms and they could also satisfy the real-time performance of AR systems.

Although the precision increases of each combined descriptor, the recall decreases. In other words, the number of correctly matched keypoints decreases, which means we detected fewer correct matched keypoints when using our combined descriptor. The reason of this may be that when two descriptors are combined into a new one, the filtering of keypoints is stricter than before. This may cause an excessive filtering on these keypoints candidates. In our future research, we will further improve this combined descriptor to deal with this over-filtering problem. Meanwhile, we will as well focus on the improvement in computation of these related algorithms to increase the speed of a real-time AR system. In the AR system, feature detection and matching are important procedures, and the light changes will affect the reliability of these steps. By using our proposed descriptor for feature matching, the stability

of the whole AR system will be improved. With the development of 5G and other technology, the AR technology will be improved in a rapid speed in the near future.

Author Contributions: Conceptualization, Zejun Xiang and Ronghua Yang; methodology, Zejun Xiang, Ronghua Yang, Chang Deng and Mingxing Teng; software, Mingxing Teng and Chang Deng; validation, Zejun Xiang, Ronghua Yang, Chang Deng, Mingxing Teng and Mengkun She; formal analysis, Zejun Xiang, Ronghua Yang and Degui Teng; investigation, Chang Deng, Mingxing Teng and Mengkun She; resources, Zejun Xiang, Ronghua Yang, Chang Deng and Mingxing Teng; data curation, Chang Deng, Mingxing Teng and Mengkun She; writing—original draft preparation, Chang Deng and Mingxing Teng; writing—review and editing, Zejun Xiang, Ronghua Yang, Chang Deng, Degui Teng and Mengkun She; visualization, Zejun Xiang, Ronghua Yang, Chang Deng and Mingxing Teng; supervision, Zejun Xiang, Ronghua Yang and Degui Teng; project administration, Zejun Xiang and Ronghua Yang; funding acquisition, Zejun Xiang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Chongqing Natural Science Foundation, grant number No.cstc2019jcyj-msxmX0153, the 111 projects, grant number No. B18062, the National Natural Science Foundation of China, grant number No. 41304001 and No. 41674005, Chongqing Natural Science Foundation and Technology Innovation Special Project of Social Undertaking and People's Livelihood Guarantee, grant number No. cstc2016shmszx0299.

Acknowledgments: The test data in this research is provided by Mikolajczyk K, and we thank all anonymous reviewers for their valuable, constructive and prompt comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yuen, S.C.-Y.; Yaoyuneyong, G.; Johnson, E. Augmented Reality: An Overview and Five Directions for AR in Education. *J. Educ. Technol. Dev. Exch.* **2011**, *4*, 119–140. [CrossRef]
2. Cheng, J.; Chen, K.Y.; Chen, W. Comparison of marker-based AR and markerless AR: A case study on indoor decoration system. *Lean Comput. Constr. Congr.* **2017**, *2*, 483–490.
3. Mahfoud, H.; Zenati, N.; Hayet, B.; Kamila, G.; Nouara, A. Object recognition based on ORB descriptor for markerless augmented reality. In Proceedings of the 9ème Conférence sur le Génie Electrique, EMP, Bordj El Bahri, Alger, 14–15 April 2015.
4. Peng, B.Y.; Wang, W.; Zhou, C.; Liu, X.W. SUSAN-SURF fast matching algorithm for augmented reality. *Appl. Res. Comput.* **2015**, *8*, 304–308.
5. Van Oosterom, P.J.M.; Quak, C.W.; Tijssen, T.P.M.; Verbree, E. The Architecture of the Geo-Information Infrastructure. In Proceedings of the UDMS 2000, 22nd Urban Data Management Symposium, Delft, The Netherlands, 11–15 September 2000; pp. 9–18.
6. Wang, Y.; Lv, M.; Meng, X.F. Image matching methods with illumination robustness. *J. Comput. Appl.* **2019**, *39*, 262–266.
7. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
8. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]
9. Tola, E.; Lepetit, V.; Fua, P. A fast local descriptor for dense matching. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.
10. Rosten, E.; Drummond, T. Machine Learning for High-Speed Corner Detection. *Appl. Evol. Comput.* **2006**, *3951*, 430–443.
11. Leutenegger, S.; Chli, M.; Siegwart, R. BRISK: Binary Robust invariant scalable keypoints. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2548–2555.
12. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. BRIEF: Binary Robust Independent Elementary Features. *Appl. Evol. Comput.* **2010**, *6314*, 778–792.
13. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 IEEE International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2012; pp. 2564–2571.
14. Alahi, A.; Ortiz, R.; Vandergheynst, P. FREAK: Fast retina keypoint. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 510–517.

15. Alcantarilla, P.F.; Bartoli, A.; Davison A., J. KAZE features. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 214–227.
16. Weickert, J.; Romeny, B.; Viergever, M. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Process.* **1998**, *7*, 398–410. [CrossRef]
17. Alcantarilla, P.F.; Solutions, T. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell.* **2011**, *34*, 1281–1298.
18. Grewenig, S.; Weickert, J.; Bruhn, A. From box filtering to fast explicit diffusion. In *Dagm Conference on Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 533–542.
19. Gevrekci, M.; Gunturk, B.K. Illumination robust interest point detection. *Comput. Vis. Image Underst.* **2009**, *113*, 565–571. [CrossRef]
20. Liu, H.; Hao, K.R.; Ding, Y.S.; Mao, M. An image matching method of new illumination-robust and anti-blur combined moment invariants. *Chin. J. Sens. Actuators* **2013**, *26*, 1258–1264.
21. Ouyang, P.; Yin, S.Y.; Liu, L.B.; Wei, S. A fast extreme illumination robust feature in affine space. In Proceedings of the 2014 International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 2365–2370.
22. Lyu, W.X.; Peng, Q.M.; Lyu, Y.Z. Illumination-robust image matching based on non-linear correlation. *J. Comput. Aided Des. Comput. Graph.* **2009**, *21*, 825–830.
23. Zhao, T.Y.; Wang, J. Improved KAZE algorithm for illumination change image matching. *Comput. Meas. Control* **2017**, *25*, 185–188.
24. Heikkilä, M.; Pietikäinen, M.; Schmid, C. Description of interest regions with local binary patterns. *Pattern Recognit.* **2009**, *42*, 425–436. [CrossRef]
25. Ojala, T.; Pietikäinen, M.; Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit. J. Pattern Recognit. Soc.* **1996**, *29*, 51–59. [CrossRef]
26. Zheng, Y.B.; Huang, X.S.; Feng, S.J. An image matching algorithm based on combination of SIFT and the rotation invariant LBP. *J. Comput.-Aided Des. Comput. Graph.* **2010**, *2*, 104–110.
27. Mikolajczyk, K.; Tuytelaars, T.; Schmid, C.; Zisserman, A.; Matas, J.; Schaffalitzky, F.; Kadir, T.; Van Gool, L. A Comparison of Affine Region Detectors. *Int. J. Comput. Vis.* **2005**, *65*, 43–72. [CrossRef]
28. Mikolajczyk, K.; Schmid, C. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1615–1630. [CrossRef] [PubMed]