LETTER

CBRISK: Colored Binary Robust Invariant Scalable Keypoints

Huiyun JING†, **Xin HE**†, **Qi HAN**†, ***Nonmembers***, ***and*** **Xiamu NIU**†a), ***Member***

SUMMARY BRISK (Binary Robust Invariant Scalable Keypoints) works dramatically faster than well-established algorithms (SIFT and SURF) while maintaining matching performance. However BRISK relies on intensity, color information in the image is ignored. In view of the importance of color information in vision applications, we propose CBRISK, a novel method for taking into account color information during keypoint detection and description. Instead of grayscale intensity image, the proposed approach detects keypoints in the photometric invariant color space. On the basis of binary intensity BRISK (original BRISK) descriptor, the proposed approach embeds binary invariant color presentation in the CBRISK descriptors. Experimental results show that CBRISK is more discriminative and robust than BRISK with respect to photometric variation.

key words: *color space, photometric invariant, binary descriptor*

1. Introduction

Feature matching is at the base of many Computer Vision applications, e.g. image retrieval, object recognition and matching, visual mapping. Due to ever growing consumption of visual content and real-time requirement for such visual applications, there exists an urgent need for fast and efficient methods to compute local descriptors and match them afterwards. It is difficult to satisfy this demand using existing methods (SIFT [1], SURF [2]) relying on costly descriptors for detection and matching.

To address the above mentioned problem, binary descriptors, the bit strings directly computed from image patches, are proposed. The similarity between the binary descriptors can be evaluated using the Hamming distance, which can be computed much faster than the Euclidean one between floating-point descriptors [1], [2] on modern CPUs.

Calonder et al. [3] proposed efficient BRIEF (Binary Robust Independent Elementary Features) descriptors that are computed using simple intensity comparisons at random pre-determined pixel locations. The performance of BRIEF is similar to SIFT in many respects, including robustness to lighting, blur, and perspective distortion. However, it is very sensitive to image rotation and scale changes restricting its application to general tasks. Rublee et al. [4] presented the oriented BRIEF descriptor ORB (Oriented FAST and Rotated BRIEF). Inspired by the rapid computation of FAST (Features from Accelerated Segment Test) corner detector [5], they applied FAST to detect keypoints and added a fast and accurate orientation operator to FAST for computing the orientation of keypoints. Then they steered BRIEF descriptor according to the orientation of keypoints for obtaining the oriented BRIEF descriptors. Although ORB is rotation invariant, it still cannot efficiently handle the scenes with scale changes. Leutenegger et al. [6] introduced a rotation and scale invariant binary descriptor BRISK (Binary Robust Invariant Scalable Keypoints). They proposed a novel scale-space FAST-based detector and generated the descriptor by comparing the intensities retrieved by dedicated sampling of each keypoint neighborhood. BRISK achieves comparable quality performance with well-established algorithms (SIFT and SURF) at much less computation time. However, BRIEF, ORB, and BRISK are all designed mainly for grayscale images.

In this letter, we aim to generate high-performance and low-computation binary descriptor for color images. We present a novel Colored BRISK (CBRISK), which not only contains the color information but also possesses good robustness with respect to color, rotation and scale changes.

The rest of the letter is organized as follows. The analysis of BRISK method is described in Sect. 2. Section 3 introduces the details of the proposed method CBRISK. Section 4 is devoted to the results and comparative analysis. Finally, the conclusion is drawn in Sect. 5.

2. Limitations of BRISK

For the sake of efficiency, BRISK applies AGAST [7] to detect keypoints, which is an extension for accelerated performance of FAST corner detector. In order to achieve invariance to scale, the interest points are identified across scale-space pyramid of the input image using the FAST score. However BRISK use the grayscale intensity image of the truecolor image as the input to detect keypoint, color invariance is not adequately considered during keypoints detection.

BRISK identifies the feature characteristic direction of each keypoint by computing the gradients of certain point pairs sampled according to a pattern around keypoint. After rotating the sampling pattern, the rotation- and scale-normalized binary descriptor is obtained by performing intensity comparison between pair-wise points inside the pattern. Nevertheless, during the generation of BRISK descriptor, only the intensity information is taken into account. Color, an important component for distinguishing color im-

Manuscript received August 6, 2012.
Manuscript revised October 23, 2012.
†The authors are with the Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China.
a) E-mail: xm.niu@hit.edu.cn
DOI: 10.1587/transinf.E96.D.392

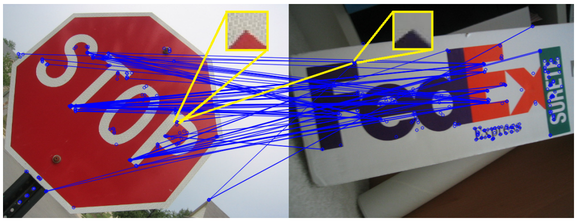

Fig. 1 BRISK descriptors without color information cause mismatch of different feature points with similar local pattern. Two practical images, from UKBench dataset [8], are matched. Blue lines connect the matched BRISK feature points. For a clear showing of mismatch, one pair of mismatched BRISK feature points are amplified, and the connected line of them is labeled yellow.

ages, is neglected. Only intensity-based binary descriptors will causes mismatch of different feature points with similar local pattern in color images (Fig. 1). So, including color information in descriptor is important for matching color images.

To solve the above-mentioned two problems, color invariant is combined with BRISK feature detector for detection of scale and color invariant features, and color information with robustness to illumination changes is added into the feature descriptor to enhance the distinctive power of BRISK. We name the proposed method "CBRISK" (Colored BRISK).

3. Colored Binary Robust Invariant Scalable Keypoints (CBRISK)

In this section, we describe the two stages in CBRISK, keypoint detection and descriptor composition. Since more stable keypoints mean better performance [9], CBRISK aims at to detect keypoints invariant to illumination and scale changes. For each keypoint, a colored local invariant binary feature descriptor is built to distinctively characterize the local region around the keypoint.

3.1 Keypoint Detection

For scale and color invariant keypoint detection, the photometric invariants proposed by Geusebroek et al. [10] are combined with BRISK feature detector. For the sake of completeness, we give a brief description of these invariants.

The color invariants from [10] are derived from the Gaussian opponent color model. In the opponent model, there are intensity, yellow-blue and red-green three channels, which are computed from RGB values directly by the linear transformation

$$\begin{bmatrix} \widehat{E}(x,y) \\ \widehat{E}_\lambda(x,y) \\ \widehat{E}_{\lambda\lambda}(x,y) \end{bmatrix} = \begin{pmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{pmatrix} \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix} \tag{1}$$

where $\widehat{E}$, $\widehat{E}_\lambda$, and $\widehat{E}_{\lambda\lambda}$ respectively denote the intensity, blue-yellow and green-red channel.

By dividing $\widehat{E}_\lambda$ by $\widehat{E}_{\lambda\lambda}$, the color invariant H is got

$$H = \frac{\widehat{E}_\lambda}{\widehat{E}_{\lambda\lambda}} \tag{2}$$

The color invariants C_λ and $C_{\lambda\lambda}$ are obtained as

$$C_\lambda = \frac{\widehat{E}_\lambda}{\widehat{E}} \qquad C_{\lambda\lambda} = \frac{\widehat{E}_{\lambda\lambda}}{\widehat{E}} \tag{3}$$

In [10], it is shown that H, C_λ, and $C_{\lambda\lambda}$ are all object reflectance properties independent of the viewpoint, surface orientation, illumination direction and illumination intensity. Because H flattens out all patterns in an image [11], which is not a desired property for detecting corner as BRISK does during keypoints detection. We hence focus on C photometric invariant.

The photometric invariant $C_{\lambda\lambda}$ is used as the working space for the input image. Then, the scale and color invariant features are detected by the BRISK feature detector AGAST at the extrema of a scale-space pyramid of the $C_{\lambda\lambda}$ invariant plane.

3.2 Descriptor Building

After detecting the keypoints, binary feature descriptors are built to describe these points. For distinguishing keypoints extracted from different images, these descriptors should contain the necessary distinct information from their corresponding keypoints [10]. It has been demonstrated that color has high discriminative power for recognizing different images [12]. However, color is sensitive to illumination and recording conditions changes, which makes the task of generating robust colored feature descriptor difficult.

In order to contain color information while preserving color invariance, several colored floating-point descriptors [9], [11], [12] based on photometric invariant have been proposed. Abdel-Hakim et al. [9] firstly made use of color invariant H as input for SIFT to generate colored SIFT descriptors. While Burghouts et at. [12] applied C invariant to generate C-SIFT descriptor and demonstrated that C-SIFT outperforms other photometric invariant based SIFT descriptors. Chu et al. [11] obtained the C-SURF descriptor according to the way producing C-SIFT, getting descriptor from the C color space consisting one intensity channel and 2 channels $\{C_\lambda, C_{\lambda\lambda}\}$.

Though including photometric invariants in the C-SIFT and C-SURF descriptors generation increases discriminative power, it is not suitable for binary descriptor BRISK, each bit of which is obtained by performing intensity comparison. Due to the nonlinear transformations used to compute photometric invariants, these photometric invariants are unstable in certain areas of the RGB-cube [13]. These instabilities makes the magnitude relationship between pair-wise photometric invariants sampled around each keypoint sensitive to color variations caused by changes in illumination, blurring, etc. It causes photometric invariants are not suited for BRISK descriptor. Furthermore, reduplicative descriptor extractions from several channels (intensity and photometric invariants channels) are necessary for producing C-SIFT and C-SURF descriptors.

Van De Sande et al. [14] pointed out that illumination changes can be modeled as a diagonal-offset model.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix} \tag{4}$$

Based on the analysis of diagonal-offset model, Van De Sande et al. [14] proposed five types of common illumination changes, light intensity change ($a = b = c$ and $o_1 = o_2 = o_3 = 0$), light intensity shift ($a = b = c = 1$ and $o_1 = o_2 = o_3$), light intensity changes and shift ($a = b = c$ and $o_1 = o_2 = o_3$), light color change ($a \neq b \neq c$ and $o_1 = o_2 = o_3 = 0$), and light color change and shift ($a \neq b \neq c$ and $o_1 \neq o_2 \neq o_3$).

The magnitude relationship of *RGB* values of each image pixel is invariant to light intensity change and shift, part light color change and shift. It inspires us to take into account the magnitude relationship of *RGB* values in CBRISK descriptor generation. We apply the magnitude relationship of *RGB* values of each pixel, sampled according to the oriented BRISK sampling pattern, to represent the local color information for each keypoint. For acquiring the bit-string expression of color information, six bits are used to represent the magnitude relationship of *RGB* values of each sampled point. The six bits are divided into two groups, each of which contains three bits respectively corresponding to each *RGB* channel. The first group represents which channel contains the maximum value, while the second group represents the channel with the minimum value. When the maximum channel (or minimum channel) is determined, the corresponding bit is set to 1, other two bits are set to 0. The distance between color bit-string representations can be directly computed by *XOR* operation without decoding these representations. After concatenating the intensity BRISK (original BRISK) descriptor with the bit-string representation of color, we obtain the CBRISK descriptor, which is invariant to rotation and color variations. During the generation of BRISK descriptor, there are no reduplicative BRISK descriptor extractions from several channels.

4. Experimental Results

The experiments were performed following the widely adopted evaluation method on the publicly available datasets†, both proposed by Mikolajczyk et at. [15]. Each of the datasets contains six images with geometric and photometric transformations, and also provides the ground truth transformations against the first image. All comparisons are made by matching the first image of each dataset against any one of the others in the same dataset. The transformations performed on these datasets include image blur (Trees and Bikes), illumination changes (Leuven), zoom and rotation (Bark), and view-point change (Wall and Graffiti).

For comparison purpose, the method C-SIFT [12] with

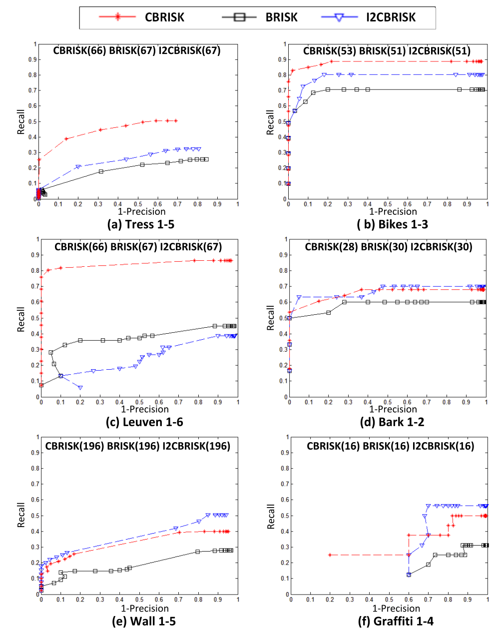

Fig. 2 Comparison results showing precision-recall curves for BRISK, CBRISK and I2CBRISK under approximately equal number of correspondences. The number of correspondences are denoted in the figures.

the best performance among the photometric invariant based colored SIFT descriptors is considerd. We extended it to BRISK, and got the new descriptors named *I2CBRISK*, which contains the descriptors extracted from $\{I, C_\lambda, C_{\lambda\lambda}\}$ channels. Then we compared the joint performance of all detection, extraction and matching stages, computational speed and the length of bit-string descriptors of *CBRISK* with *BRISK* (original implementation††) and *I2CBRISK*.

4.1 Evaluation and Comparison of CBRISK

Recall and 1-precision [15] are used as the evaluation criterions. The evaluation criterions were computed using the MATLAB evaluation scripts provided by Mikolajczyk et al.†††. And the nearest neighbor matching strategy is chosen to gain the numbers of matches and correct matches, which finds the nearest pair of keypoints.

We evaluated the performance of *CBRISK* and compared it to *BRISK* and *I2CBRISK*. Figure 2 shows the the precision-recall curves for a selection of image pairs of different datasets. For the sake of fairness, we adapted the detection thresholds leading to an approximately equal number of correspondences for *CBRISK*, *BRISK* and *I2CBRISK*.

As illustrated in Fig. 2, *CBRISK* outperforms *BRISK* in all datasets. In image blur (Fig. 2 (a) and (b)) and illumination change (Fig. 2 (c)) cases, *CBRISK* get better performance than *I2CBRISK*. *BRISK* performs competitively with *I2CBRISK* in other cases. The performance of

†http://www.robots.ox.ac.uk/~vgg/research/affine/
††http://www.asl.ethz.ch/people/lestefan/personal/BRISK
†††http://www.robots.ox.ac.uk/~vgg/research/affine/desc_evaluation.html

Table 1 Performance comparison.

(a) Description timings for the third image in the Wall sequence

	BRISK	**I2CBRISK**	**CBRISK**
Detection threshold	80	80	60
Number of points	254	254	259
Description time [ms]	37	105	56
Time per point [ms]	0.1456	0.4134	0.2162

(b) Length of bit-string descriptors

	BRISK	**I2CBRISK**	**CBRISK**
bits	512	1536	872

I2CBRISK in the Trees dataset is reduced, since the magnitude relationship between pair-wise color invariants is sensitive to large illumination change.

4.2 Timings

We used *CBRISK*, *BRISK* and *I2CBRISK* to generate local feature descriptors for the images in the datasets. Table 1 (a) compares the average description times on the third image of the Wall sequence. For *BRISK*, we used the authors' implementations. *CBRISK* and *I2CBRISK* are implemented in C++ based on *BRISK* code. Timings were recorded on a laptop with a dual-core $i5$ 2.67 GHz processor (only using one core) running Window7 (32-bit).

Since *I2CBRISK* needs reduplicative descriptor extractions for one intensity channel and 2 color invariant channels $\{C_\lambda, C_{\lambda\lambda}\}$, the description timing of *I2CBRISK* is approximately three times of *BRISK*. However, *CBRISK* simply compares the R G B values of each pixel sampled around the keypoint while extracting *BRISK* descriptor, the description timing of *CBRISK* increases a half times of *BRISK*.

4.3 Descriptor Lengths

Table 1 (b) shows the lenghts of *BRISK*, *I2CBRISK* and *CBRISK* descriptors. The bit-string of *BRISK* descriptor contains 512 bits. Dut to extracting *BRISK* descriptor for three channels, the length of *I2CBRISK* is three times of *I2CBRISK*. Since the sampling pattern of *BRISK* contains 60 points around a interest point, *CBRISK* obtains a bit-string of length 360 for representing the magnitude relationship of *RGB* values for 60 points. Combining the original *BRISK* descriptor and the bit-string expression of color, *CBRISK* descriptor contains 872 bits.

5. Conclusions

In this paper, we introduced a novel method named CBRISK, which tackles the problem of the intensity-based BRISK lacking photometric invariance and discriminative power. In keypoint detection, the photometric invariant is used as the input to construct a scale-space pyramid, where the corners are identified as keypoints by CBRISK. For a colored invariant descriptor, the magnitude relationship of RGB values of each pixel sampled arround the keypoint is applied as the invariant color information. Then the CBRISK feature descriptor is gained by combining both the binary presentation of invariant color information and the original binary intensity BRISK descriptor. The experiments show CBRISK outperforms BRISK and I2CBRISK generated according to the C-SIFT method.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (60832010, 61100187), the Fundamental Research Funds for the Central Universities (Grant No. HIT. NSRIF. 2010046).

References

[1] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, no.2, pp.91–110, 2004.
[2] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," Computer Vision–ECCV 2006, pp.404–417, 2006.
[3] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: binary robust independent elementary features," European Conference on Computer Vision, pp.778–792, 2010.
[4] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," IEEE International Conference on Computer Vision, pp.2564–2571, 2011.
[5] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," European Conference on Computer Vision, pp.430–443, 2006.
[6] S. Leutenegger, M. Chli, and R. Siegwart, "Brisk: Binary robust invariant scalable keypoints," IEEE International Conference on Computer Vision, pp.2548–2555, 2011.
[7] E. Mair, G. Hager, D. Burschka, M. Suppa, and G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," European Conference on Computer Vision, pp.183–196, 2010.
[8] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.2161–2168, 2006.
[9] A. Abdel-Hakim and A. Farag, "Csift: A sift descriptor with color invariant characteristics," IEEE Conference on Computer Vision and Pattern Recognition, pp.1978–1983, 2006.
[10] J. Geusebroek, R. van den Boomgaard, A. Smeulders, and H. Geerts, "Color invariance," IEEE Trans. Pattern Anal. Mach. Intell., vol.23, no.12, pp.1338–1350, 2001.
[11] D. Chu and A. Smeulders, "Color invariant surf in discriminative object tracking," ECCV Workshop on Color and Reflectance in Imaging and Computer Vision, 2010.
[12] G. Burghouts and J. Geusebroek, "Performance evaluation of local colour invariants," Comput. Vis. Image Understand., vol.113, no.1, pp.48–62, 2009.
[13] J. Van De Weijer, T. Gevers, and J. Geusebroek, "Edge and corner detection by photometric quasi-invariants," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.4, pp.625–630, 2005.
[14] K. Van De Sande, T. Gevers, and C. Snoek, "Evaluating color descriptors for object and scene recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.9, pp.1582–1596, 2010.
[15] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.10, pp.1615–1630, 2005.