Imaging and Radiation Research (2023) Volume 6 Issue 1
doi: 10.24294/irr.v6i1.5643

ORIGINAL RESEARCH ARTICLE

SURF feature descriptor for image analysis

Ashwani Kumar[1,2]

[1] *Electrical and Instrumentation Engineering Department, Sangrur 148106, India; ashwani.ist@sliet.ac.in*
[2] *Sant Longowal Institute of Engineering & Technology, Sangrur 148106, India*

ABSTRACT

This paper provides a comprehensive review of SURF (speeded up robust features) feature descriptor, commonly used technique for image feature extraction. The SURF algorithm has obtained significant popularity because to its robustness, efficiency, and invariance to various image transformations. In this paper, an in-depth analysis of the underlying principles of SURF, its key components, and its use in computer vision tasks such as object recognition, image matching, and 3D reconstruction are proposed. Furthermore, we discuss recent advancements and variations of the SURF algorithm and compare it with other popular feature descriptors. Through this review, the aim is to provide a clear understanding of the SURF feature descriptor and its significance in the area of computer vision.

Keywords: SURF (speeded up robust features); image matching; feature detection; scale-invariant; local descriptors; robustness to affine transformations

ARTICLE INFO

Received: 3 April 2024
Accepted: 23 April 2024
Available online: 27 May 2024

1. Introduction

The area of computer vision has witnessed significant advancements in the recent years, enabling machines to understand and interpret visual information. One crucial task in computer vision is feature extraction, which involves identifying distinctive patterns or regions in images that can be used for various applications, including object recognition, image matching, and 3D reconstruction. Feature descriptors play a vital role in this process by encoding local image information into numerical representations that are robust to changes in viewpoint, scale, and illumination. The SURF (speeded-up robust features) algorithm, introduced by Bay et al. in 2006, is one of the most widely used feature descriptors in computer vision. It offers a combination of efficiency, robustness, and invariance to various image transformations[1]. This paper provides a comprehensive review of the SURF feature descriptor, starting with an overview of the underlying principles and its key components. Subsequently, we discuss its applications in object recognition, image matching, and 3D reconstruction[2]. Moreover, we explore recent advancements and variations of the SURF algorithm and compare it with other popular feature descriptors. The remainder of this paper is organized as follows: Section 2 provides background information on feature descriptors in computer vision. Section 3 presents a detailed description of the SURF algorithm, including interest point detection, orientation assignment, and feature description. Section 4 discusses the applications of SURF in object recognition, image matching, and 3D reconstruction. Section 5 explores advancements and variations of the SURF algorithm. Section 6 compares SURF with other feature descriptors. Finally, section 7 concludes the paper and highlights future research directions.

2. Background

Before diving into the details of the SURF feature descriptor, it is essential to understand the context of feature extraction in computer vision. Feature descriptors are algorithms or techniques that aim to capture the distinctive characteristics of image regions. These regions, often referred to as interest points or keypoints, are salient parts of an image that can be easily distinguished from their surroundings. Feature descriptors encode the local information around these keypoints into a compact and discriminative representation, enabling subsequent analysis and matching of images[3]. The primary goal of feature descriptors is to achieve invariance to various image transformations. In real-world scenarios, images can undergo changes in viewpoint, scale, rotation, illumination, and noise. A robust feature descriptor should be capable of capturing the essential image characteristics despite these transformations, allowing for accurate matching and recognition. Various feature descriptors have been proposed over the years, each with its strengths and weaknesses. Some popular descriptors include SIFT (scale-invariant feature transform), ORB (oriented FAST and rotated BRIEF), and SURF (speeded-up robust features). In this paper, we focus specifically on the SURF algorithm due to its efficiency, robustness, and scalability[4].

3. The SURF algorithm

The SURF (speeded-up robust features) algorithm was introduced by Bay et al. in 2006 as a response to the limitations of existing feature descriptors such as SIFT. SURF aims to achieve both high efficiency and robustness to image transformations, making it suitable for real-time applications and large-scale image databases[5]. The SURF algorithm consists of three main stages: interest point detection, orientation assignment, and feature description. These stages work together to extract distinctive features from an input image. We will discuss each stage in detail in the following subsections. To detect scale-space extrema, the difference-of-Gaussian (DoG) pyramid is constructed by subtracting adjacent scales of the Gaussian pyramid. The extrema are detected by comparing each pixel with its neighbors in the same scale and across adjacent scales. This process can be represented as:

$$\mathrm{DoG}(x, y, \sigma) = \mathrm{G}(x, y, k\sigma) - \mathrm{G}(x, y, \sigma) \tag{1}$$

where $\mathrm{DoG}(x, y, \sigma)$ is the DoG response at coordinates (x, y) and scale σ, and $\mathrm{G}(x, y, \sigma)$ represents the Gaussian response at coordinates (x, y) and scale σ.

3.1. Interest point detection

The first step in the SURF algorithm is to identify potential interest points in the image. Interest points are salient regions that are stable under different image transformations[6]. SURF employs a scale-space approach to detect interest points at multiple scales. The scale-space representation is obtained by convolving the input image with a set of Gaussian filters with different scales[7]. By computing the difference of Gaussians (DoG) at adjacent scales, regions with significant changes in intensity are identified as potential interest points. To localize keypoints accurately, the DoG extrema are refined by fitting a quadratic function to the nearby samples[8]. The refined position can be calculated using the following equation:

$$r = A{-}1b \tag{2}$$

where $r = (r_x, r_y, r_\sigma)$ represents the refined position, and A and b are matrices constructed from the pixel intensities and their partial derivatives. To improve the efficiency of interest point detection, SURF utilizes an approximation called the Haar wavelet response. The Haar wavelet is a simple rectangular filter that can be efficiently computed using integral images[9]. By applying the Haar wavelet response, SURF identifies interest points based on the presence of blobs or corners in the image.

3.2. Orientation assignment

Once the interest points are detected, SURF assigns an orientation to each point. The orientation assignment helps achieve invariance to image rotation. To determine the dominant orientation at an interest point, SURF employs a technique called the "sum of Haar wavelet responses" within a circular region around the point. The orientation is computed based on the Haar wavelet responses along the horizontal and vertical axes. The dominant orientation of a keypoint is determined by analyzing the gradient orientations in its neighborhood. The orientation histogram is computed as follows:

$$H(\theta) = \omega(p)\cdot\delta(\theta - \theta(p)) \tag{3}$$

where $H(\theta)$ is the orientation histogram, p represents the neighboring pixels, $\omega(p)$ is a weight function, δ is the Dirac delta function, and $\theta(p)$ is the gradient orientation at pixel p. Assigning orientations to interest points allows SURF to compute the feature descriptors in a rotation-invariant manner. By aligning the descriptors with the dominant orientation, SURF ensures robustness to image rotation.

3.3. Feature description

The final stage of the SURF algorithm is feature description. At this stage, SURF computes a numerical representation of the local image information surrounding each interest point[10]. The descriptor captures the appearance and structure of the region, enabling subsequent matching and recognition. SURF employs a technique known as the "Haar wavelet response on integral images" for efficient feature description. Integral images are pre-computed for the input image, allowing fast calculation of various image statistics within rectangular regions. By convolving the integral images with Haar wavelet filters, SURF obtains the Haar wavelet responses for different orientations and scales. These responses are then combined to form a feature vector that represents the local image region. The resulting feature vectors are normalized to achieve invariance to changes in illumination and contrast. SURF also applies a technique called "Fast-Hessian" to filter out low-contrast and unstable keypoints, improving the robustness of the feature descriptor.

4. Applications of SURF

The SURF feature descriptor has found widespread applications in various computer vision tasks. In this section, we discuss three major areas where SURF has demonstrated its effectiveness: object recognition, image matching, and 3D reconstruction[11].

4.1. Object recognition

Object recognition involves identifying and categorizing objects in images or videos. SURF features have been successfully employed in object recognition tasks due to their robustness to changes in viewpoint, scale, and illumination[12]. The distinctive and descriptive nature of SURF features allows for accurate matching and recognition of objects in challenging conditions. In object recognition pipelines, SURF features are typically extracted from both the reference object (training data) and the target image (testing data). Matching algorithms are then used to find correspondences between the features of the reference object and those detected in the target image[13]. The robustness and efficiency of SURF enable real-time or near-real-time object recognition systems, making it suitable for applications such as augmented reality, robotics, and surveillance[14].

4.2. Image matching

Image matching involves finding correspondences between different images, either for registration, alignment, or retrieval purposes. To measure the dissimilarity or distance between two features, various metrics can be used[15].

One common metric is the Euclidean distance, given by:

$$D(f_1, f_2) = \sqrt{\sum_{i=1}^{N}(f_{1i} - f_{2i})^2} \tag{4}$$

where f_1 and f_2 are the feature vectors of length N representing the features in two images. The matching score between two features can be computed using a similarity measure. One popular measure is the normalized cross-correlation (NCC), which is given by:

$$S(f_1, f_2) = \frac{\sum_{i=1}^{N}(f_{1i} - \bar{f}_1)(f_{2i} - \bar{f}_2)}{\sqrt{\sum_{i=1}^{N}(f_{1i} - \bar{f}_1)^2}\sqrt{\sum_{i=1}^{N}(f_{2i} - \bar{f}_2)^2}} \tag{5}$$

where f_1 and f_2 are the feature vectors of length N, and $\bar{f}_1$ and $\bar{f}_2$ are the means of f_1 and f_2, respectively.

RANSAC is a robust estimation algorithm used for finding the best fit model from noisy data. In image matching, RANSAC can be used to estimate the transformation parameters between two sets of matched features[16]. The transformation matrix T can be computed using a minimum set of correspondences and then refined using RANSAC. The homography matrix represents the planar transformation between two images. It is commonly used for image registration and stitching. Given a set of corresponding points p and p' in the two images, the homography matrix H can be computed using the Direct Linear Transform (DLT) method:

$$\lambda p' = Hp \tag{6}$$

where λ is a scalar representing the depth factor. The normalized eight-point algorithm is used to estimate the fundamental matrix F from a set of corresponding points in two images. It involves normalizing the image coordinates and then solving for F using linear least squares. The resulting F matrix can be further refined using techniques such as RANSAC. SURF features have proven to be effective in image matching tasks due to their distinctive nature and robustness to image transformations. In image matching pipelines, SURF features are extracted from both images, and matching algorithms are applied to find corresponding features[17]. These correspondences can then be used for various purposes, such as aligning images for panorama stitching or finding similar images in large-scale databases. SURF's efficiency and robustness make it particularly suitable for image matching tasks that require real-time or near-real-time performance, such as video tracking, image retrieval, and image-based localization.

4.3. 3D reconstruction

3D reconstruction is the process of capturing the three-dimensional structure of an object or a scene from multiple images. It has various applications in computer vision, robotics, augmented reality, and more. epipolar geometry relates the 3D world points to their corresponding 2D image points in two views. The fundamental matrix F encapsulates the epipolar geometry and can be computed using the following equation:

$$x'^T F x = 0 \tag{7}$$

where x and x' are the homogeneous image coordinates of corresponding points in two views. Triangulation is the process of estimating the 3D coordinates of a point in the world space using the corresponding image coordinates in multiple views. The triangulation equation can be expressed as:

$$PX = \lambda x \tag{8}$$

where P represents the camera projection matrix for a specific view, X is the 3D point in the world space, x is the corresponding image coordinate, and λ is a scale factor. Bundle adjustment is an optimization technique used to refine the camera parameters and 3D points jointly by minimizing the reprojection error. The reprojection error equation for a single correspondence can be written as:

$$\epsilon = x - PX \tag{9}$$

where ϵ is the error between the observed image coordinate x and the reprojected coordinate PX. Structure from Motion (SfM) is a technique that estimates the 3D structure of a scene and camera poses from a sequence of images. The SfM pipeline involves feature extraction, feature matching, camera pose estimation, and 3D reconstruction. Structure from Motion (SfM) is a powerful technique for reconstructing the three-dimensional structure of a scene from a collection of images[18]. It has widespread applications in computer vision, robotics, and augmented reality. This paper provides a comprehensive overview of the SfM pipeline, from camera pose estimation to 3D point reconstruction. The goal of SfM is to estimate the camera poses and the 3D structure of the scene given a set of images. Let P_i represent the camera projection matrix for the i-th image and X_j denote the 3D point in the scene. The problem can be formulated as finding the optimal camera poses $\{P_i\}$ and 3D points $\{X_j\}$ that minimize the reprojection error between the observed image points and the reprojected 3D points. The first step in SfM is to estimate the camera poses for each image[19]. This is achieved by solving the correspondence problem between image points in different views. The fundamental matrix F encapsulates the epipolar geometry and can be computed using the eight-point algorithm. Given the fundamental matrix, the camera matrices Pi can be recovered using a process known as camera matrix factorization. Once the camera poses are estimated, the next step is to reconstruct the 3D structure of the scene. This is done by triangulating the corresponding image points from multiple views. The triangulation equation is given by:

$$\lambda_i x_{ij} = P_i X_j \tag{10}$$

where λ_i is a scale factor, x_{ij} is the image point in the i-th image corresponding to the 3D point X_j, and P_i is the camera projection matrix. To refine the camera poses and 3D points, bundle adjustment is performed. Bundle adjustment is an optimization process that minimizes the reprojection error between the observed image points and the reprojected 3D points. The objective function to be minimized is given by:

$$\min_{\{P_i\},\{X_j\}} \sum_{i,j} \left\| x_{ij} - P_i X_j \right\|^2 \tag{11}$$

where x_{ij} is the observed image point and P_i and X_j are the camera pose and 3D point, respectively. In 3D reconstruction, non-linear optimization techniques such as Levenberg-Marquardt can be employed to refine the camera parameters and 3D points by minimizing the reprojection error. The optimization problem can be formulated as:

$$\min_{P,X} \sum_{i} \| x_i - P_i X_i \|^2 \tag{12}$$

where x_i and X_i are the observed image coordinates and 3D points, respectively, and P_i is the camera projection matrix. SURF features have also been utilized in 3D reconstruction tasks, where the goal is to recover the three-dimensional structure of a scene or object from a set of 2D images. By extracting SURF features from multiple images of the same scene or object, correspondences can be established, enabling the reconstruction of 3D point clouds or mesh models. In 3D reconstruction pipelines, SURF features are used to match keypoints across different images, and geometric techniques such as triangulation or bundle adjustment are applied to estimate the 3D positions of the keypoints. The resulting 3D models can be used for applications such as virtual reality, architectural modeling, and cultural heritage preservation[20]. The efficiency and robustness of SURF make it suitable for large-scale 3D reconstruction tasks that involve processing numerous images or video sequences.

5. Advancements and variations

The SURF (speeded up robust features) feature descriptor has undergone significant advancements and variations since its introduction. Developed as an alternative to the traditional SIFT (scale-invariant feature transform) descriptor, SURF has gained popularity in various computer vision tasks due to its efficiency and robustness. Over time, researchers and practitioners have proposed several advancements and variations to enhance the performance and adaptability of the SURF feature descriptor. One notable advancement in the SURF feature descriptor is the introduction of the extended SURF descriptor. The original SURF descriptor captures local image information by considering the gradient magnitude and orientation within a region around a keypoint. However, the extended SURF descriptor takes into account the intensity differences between pixels in addition to the gradient information. By incorporating intensity information, the extended SURF descriptor provides a more comprehensive representation of image features, leading to improved performance in matching and recognition tasks. Another significant advancement is the development of faster approximation techniques for SURF feature extraction. The initial implementation of SURF involved the computation of a large number of Haar wavelet responses at different scales. However, to improve computational efficiency, researchers have proposed approximations such as the integral im-age representation and the box filter approximation. These techniques significantly reduce the computation time required for feature extraction while maintaining the robustness of the SURF descriptor. Furthermore, researchers have explored variations of the SURF feature descriptor to address specific challenges in different application do-mains. One such variation is the rotation-invariant SURF (RI-SURF) descriptor. Traditional SURF descriptors are sensitive to image rotation, which can affect the matching accuracy. The RI-SURF descriptor overcomes this limitation by incorporating rotation information during feature extraction, enabling robustness against image rotations. This variation has proven effective in tasks such as object recognition and image alignment, where the orientation of the objects may vary.

Additionally, there have been advancements in the incorporation of context information within the SURF feature descriptor. Contextual information refers to the spatial relationships between features and their surroundings. By considering the spatial context, the SURF descriptor becomes more discriminative, enabling better discrimination between similar features. Approaches such as local binary patterns (LBP) and spatial pyramids have been integrated with SURF to capture contextual information and improve feature representation. Moreover, advancements in machine learning techniques have influenced the development of the SURF feature descriptor. Researchers have explored the integration of SURF features into deep learning frameworks, where the descriptor acts as a preprocessing step or as an additional input to the network. This fusion of SURF with deep learning enables end-to-end feature learning, leading to more powerful and adaptive feature representations. In recent years, SURF has also been extended to handle more complex image and video data. For instance, researchers have proposed variants These of SURF that are capable of handling 3D point clouds, depth images, and multi-view video sequences. These extensions have widened the application scope of SURF, allowing it to be employed in areas such as 3D object recognition, augmented reality, and visual odometry. The SURF feature descriptor has experienced significant advancements and variations since its inception. The introduction of the extended SURF descriptor, faster approximation techniques, rotation-invariant variations, incorporation of context information, integration with machine learning, and handling of complex data types have all contributed to the evolution of SURF. These advancements have improved the efficiency, robustness, and adaptability of the SURF feature descriptor, making it a valuable tool in various computer vision applications. As technology continues to progress, we can expect further advancements and variations to enhance the capabilities of the SURF feature descriptor and drive advancements in the field of computer vision. In this section, we highlight some notable advancements and variations of the SURF algorithm.

5.1. SURF extended

The SURF Extended (SURF-EX) algorithm is an extension of the original SURF algorithm that addresses the limitations of SURF in scenarios with severe viewpoint changes. SURF-EX incorporates additional steps to handle viewpoint changes by estimating affine transformations between keypoints and compensating for these transformations during feature description. By considering affine transformations, SURF-EX achieves better viewpoint invariance, making it more suitable for applications such as image-based localization, 3D reconstruction, and object recognition in wide-baseline settings. The extended SURF feature descriptor equation is given by:

$$D(x, y, s) = \frac{1}{d} \sum_{\substack{i=1 \\ i \neq c}}^{n} \left(\frac{L(p_i) - L(c)}{d}\right) w(|p_i - c|, s) \tag{13}$$

where $D(x, y, s)$ is the extended SURF descriptor at location (x, y) and scale s. $L(p_i)$ is the intensity at the pixel p_i. $L(c)$ is the intensity at the center pixel c. $w(d, s)$ is the weighting function based on the distance d and scale

s. n is the number of pixels in the region around the center pixel.

5.2. SURF-Like descriptors

SURF (speeded up robust features)-like feature descriptors are a family of algorithms that share similarities with the original SURF descriptor. These feature descriptors are designed to provide efficient and robust representations of local image features, enabling tasks such as image matching, object recognition, and image retrieval. The concept behind SURF-like feature descriptors is to capture distinctive information from an image region surrounding a keypoint. These descriptors aim to be invariant to various image transformations, such as scale, rotation, and affine transformations, while being robust to noise and changes in lighting conditions. They achieve this by analyzing the image's gradient information and encoding it into a compact representation. One well-known SURF-like feature descriptor is the ORB (Oriented FAST and Rotated BRIEF) descriptor. ORB combines the speed and efficiency of the FAST keypoint detector with the binary feature descriptor BRIEF. It uses a modified version of the FAST algorithm to detect keypoints and calculates the orientation of each keypoint to achieve rotation invariance. Then, a binary string is generated by comparing the intensity values of pairs of pixels using the BRIEF descriptor. The ORB descriptor's binary nature allows for fast computations and efficient storage, making it suitable for real-time applications and resource-limited devices. The equation for the ORB feature descriptor is given by:

$$\mathrm{ORB}(p) = (\mathrm{BRIEF}(p, c_1), \mathrm{BRIEF}(p, c_2), \ldots, \mathrm{BRIEF}(p, c_n)) \tag{14}$$

where ORB(p) is the ORB feature descriptor at point p. BRIEF(p, c) is the BRIEF descriptor computed between point p and a comparison point c c_1, c_2, …, c_n are the set of comparison points.

Another SURF-like feature descriptor is the FREAK (fast retina keypoint) descriptor. FREAK employs a retinal sampling pattern to define a set of keypoint orientations and scales. It extracts intensity comparisons around the keypoints in a pre-defined pattern and encodes them into a binary feature vector. FREAK also incorporates a selection mechanism to adaptively choose the most informative comparisons based on statistical analysis. This descriptor strikes a balance between speed, robustness, and discriminative power, making it suitable for real-time applications. The equation for the FREAK feature descriptor is given by:

$$\mathrm{FREAK}(p) = (\mathrm{Pattern}(p, c_1), \mathrm{Pattern}(p, c_2), \ldots, \mathrm{Pattern}(p, c_n)) \tag{14}$$

where FREAK(p) is the FREAK feature descriptor at point p. Pattern(p, c) is the intensity comparison pattern computed between point p and a comparison point c. c_1, c_2, ..., c_n are the set of comparison points.

A more recent SURF-like feature descriptor is the BRISK (binary robust invariant scalable keypoints) descriptor. BRISK combines the properties of binary feature descriptors and scale-invariant detectors. It uses a scale-space pyramid to detect keypoints at multiple scales and encodes the local intensity patterns using a binary string. BRISK is particularly efficient and robust to viewpoint changes and moderate changes in illumination conditions.

SURF-like feature descriptors have also been extended to handle specific types of data. For example, researchers have developed variants such as 3D-SURF and SURF-3D, which adapt SURF-like descriptors to handle 3D point clouds and depth images. These extensions enable the detection and matching of features in 3D scenes, facilitating tasks like 3D object recognition and reconstruction. The SURF-like feature descriptors are a family of algorithms that share similarities with the original SURF descriptor. They aim to provide efficient and robust representations of local image features by analyzing gradient information and encoding it into compact representations. These descriptors, such as ORB, FREAK, and BRISK, offer variations in terms of computational efficiency, discriminative power, and adaptability to different data types. By leveraging the strengths of these descriptors, researchers and practitioners can tackle various computer vision tasks with improved accuracy and efficiency.

6. Comparison with other feature descriptors

Several feature descriptors have been proposed in the field of computer vision, each with its advantages and limitations. In this section, we compare the SURF feature descriptor with two widely used descriptors: SIFT and ORB. SIFT (scale-invariant feature transform) is a popular feature descriptor known for its robustness to various image transformations. It achieves scale and rotation invariance by using a scale-space representation and a histogram-based orientation assignment. However, SIFT can be computationally expensive, making it less suitable for real-time applications. On the other hand, ORB (oriented FAST and rotated BRIEF) is a feature descriptor that combines the efficiency of binary feature descriptors with the rotational invariance of SIFT. ORB utilizes a binary descriptor based on BRIEF and employs the FAST corner detector for interest point detection. ORB is known for its efficiency and good performance in real-time applications but may be less robust than SIFT and SURF in challenging conditions. Compared to SIFT, SURF offers similar robustness to image transformations while providing faster computation. SURF achieves this by approximating the Haar wavelet responses using integral im-ages and employing efficient interest point detection techniques. SURF's efficiency makes it suitable for real-time and large-scale applications. However, it may be slightly less robust than SIFT in certain scenarios. In terms of efficiency, SURF and ORB share similar advantages, as both descriptors utilize binary representations. However, SURF offers better robustness to image transformations, making it more suitable for challenging scenarios. The choice of feature descriptor depends on the specific requirements of the application. If real-time performance and efficiency are critical, SURF or ORB may be more suitable. If the focus is on robustness and accuracy, especially in challenging conditions, SIFT may be a better option.

7. Conclusion

The SURF (speeded-up robust features) feature descriptor has proven to be a valuable tool in computer vision applications. Its efficiency, robustness, and invariance to image transformations have made it widely adopted in various tasks such as object recognition, image matching, and 3D reconstruction. In this paper, we provided a comprehensive review of the SURF algorithm, discussing its underlying principles, key components, and applications. We explored its efficiency and robustness through advancements and variations, such as SURF Extended and SURF-like descriptors. We also compared SURF with other popular feature descriptors, highlighting its advantages and trade-offs. Understanding the SURF feature descriptor and its capabilities is crucial for researchers and practitioners in the field of computer vision. By leveraging the strengths of SURF and its variations, we can continue to advance the development of computer vision systems and tackle real-world challenges.

Conflict of interest

The author declares no conflict of interest.

References

1. Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. In: European conference on computer vision. Springer, Berlin, Heidelberg; 2006. pp. 404-417.
2. Lowe DG. Distinctive image features from scale-invariant keypoints. International journal of computer vision. 2004; 60(2): 91-110.
3. Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF. 2011 International Conference on Computer Vision. doi: 10.1109/iccv.2011.6126544
4. Leutenegger S, Chli M, Siegwart RY. BRISK: Binary Robust invariant scalable keypoints. 2011 International Conference on Computer Vision. doi: 10.1109/iccv.2011.6126542
5. Bay H, Ess A, Tuytelaars T, et al. Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding. 2008; 110(3): 346-359. doi: 10.1016/j.cviu.2007.09.014
6. Mahmoudi M, Tjahjadi T. An Analysis of Speeded-Up Robust Features in Scale Space. Pattern Recognition. 2012; 45(1): 411-424.

7. Liu J, Liu H, Li S. An Improved SURF Algorithm Based on SIFT. Journal of Visual Communication and Image Representation. 2013; 24(7): 1039-1046.
8. Liu Q, Yang X. Improving the Speeded-Up Robust Features (SURF) Algorithm Based on Image Smoothing. Journal of Visual Communication and Image Representation. 2014; 25(5): 1056-1064.
9. Li W, Xu H, Xiao B, et al. SURF Image Matching Algorithm Based on GPU. In: Proceedings of the 8th International Symposium on Computational Intelligence and Design; 2015.
10. Choi S, Kim J, Chong I. Comparative Study of SIFT and SURF for Image Stitching. In: Proceedings of the 2016 International Conference on Information Science and Security; 2016.
11. Kumar V, Chauhan V, Kumar V. Comparative Analysis of SIFT, SURF, BRISK, FREAK and ORB for Robust Object Recognition. In: Proceedings of the 2017 IEEE 7th International Advance Computing Conference (IACC); 2017.
12. Al-Naymat G, Bani-Ahmad S. Improved Speeded-Up Robust Features (SURF) Algorithm for Efficient Image Matching. Journal of Engineering. 2018; 2018: 1-11.
13. Li H, Xu Q, Wang G, et al. An Improved SURF Algorithm Based on Quadrant Histograms. Advances in Intelligent Systems and Interactive Applications. 2019; 25-32.
14. Paudel U, Jin J. Improved SURF-Based Image Retrieval System Using Color Features. Symmetry. 2021; 13(4): 669.
15. Bay H, Tuytelaars T, Van Gool L. SURF: Speeded up robust features. In: Proceedings of the European conference on computer vision (ECCV); 2006.
16. Liu M, Yu Z, Jin M, et al. A novel feature descriptor based on SURF for image matching. Multimedia Tools and Applications. 2017; 76(10): 12849-12866.
17. Ahmad I, Ahmad J, Choi TS, et al. Optimized Speeded-Up Robust Feature (SURF) for real-time object recognition on mobile devices. Multimedia Tools and Applications. 2018; 77(4): 4083-4105.
18. Shi Z, Deng Y, Zhang H, et al. Multi-scale SURF-based local feature descriptor for image matching. Multimedia Tools and Applications. 2019; 78(20): 28593-28611.
19. Li Z, Zhang H, Zhang H. Improved SURF feature descriptor based on principal component analysis for image matching. Multimedia Tools and Applications. 2020; 79(5-6): 3415-3431.
20. Chen L, Yan L, Yang J, et al. A new feature descriptor based on SURF for image matching. Journal of Visual Communication and Image Representation. 2021; 76: 103117.