Speeded up robust feature (SURF): Survey with application

Dr. Matheel E. Abdulmunem *, Zainab H.Fatoohi*
***University of Technology, Computer Science Department, Baghdad, Iraq**
matheel_74@yahoo.com, zainab.fatoohi@gmail.com

Abstract

The Speeded up Robust Features (SURF) strategy is a fast and strong algorithm for nearby, similitude invariant portrayal and examination of images. It can be utilized for errands, for example, object recognition, image registration, classification or 3D reconstruction. Which is incompletely enlivened by the Scale-Invariant Feature Transform (SIFT) descriptor.

It approximates or even outflanks already proposed plans concerning uniqueness, sturdiness, and being repeatable, so can be figured and analyzed substantially speedier. It utilizes the integral image, Hessian matrix and Haar wavelet responses to improve the performance in robust way. The general structure can be partitioned into three stages: interest point detection, interest point description, and interest point matching.

In this survey paper, SURF was introduced and envelops a definite portrayal of the detector and descriptor and afterward investigates the effects of the most imperative parameters, then SURF properties, advantages and drawbacks was discussed, also comparison between SIFT and SURF was made. Finally, the application of SURF on retinal image was presented.

Keyword: Speeded up Robust Feature, Scale-Invariant Feature Transform.

المستخلص

Speeded up Robust Feature (SURF) هي خوارزميه سريعه وقويه للتمثيل المحلي والتشابه الثابت وفحص الصور. ويمكن استخدامها لمهام مثل التعرف على الكائنات وتسجيل الصور والتصنيف او اعاده التشكيل ثلاثي الابعاد. وهي مستوحاة جزئيا من خوارزميه

Scale-Invariant Feature Transform(SIFT) .

وهي تتفوق على الخوارزميات المقترحه سابقا فيما يتعلق بالتميز, المتانه, وكونها قابله للتكرار, ايضا يمكن حسابها وتحليلها بشكل اسرع. كونها تستخدم الصوره المتكامله, مصفوفه ال Hessian, الاستجابات المويجيه Haar لتحسين الاداء بطريقه قويه. ويمكن تقسيم الاطار العام للعمل الى ثلاثه مراحل: الكشف عن نقاط الفائده, وصف نقاط الفائده, ومطابقه نقاط الفائده.

في هذه الورقه الاستقصائيه, تم تقديم ال SURF مع وصف تفصيلي للكاشف والواصف ومن ثم اكتشاف تأثيرات اهم المتغيرات, ثم تمت مناقشه خصائص ال SURF, ميزاتها, وعيوبها, كما اجريت مقارنه بين ال SIFT و ال SURF, واخيرا تم طرح تطبيقات ال SURF على صور شبكيه العين.

1. Introduction

The function of discovering point symmetries between two images of a similar view or object is a piece of numerous applications of computer vision. Speeded-Up Robust Features (SURF) can be depicted as a feature extraction and point relationship system for use with two-dimensional images.
The scan for separate image point symmetries has 3 principle stages. First, keypoints such as corners, blobs, and T-junctions are selected at distinctive locations in the image (detector). Next, using feature vector the neighborhood of every interest point is represented (descriptor). Finally, between various images the descriptor vectors are matched based on a distance between the vectors, e.g. the Euclidean distance [1].
The detector ought to contain the property of repeatability to seek out same physical interest points in distinct viewing conditions. The descriptor ought to contain the properties of distinctive, sturdy to noise, geometric and mensuration deformations. The matching procedure depends on the space between descriptors and this could be created quicker via utilizing less dimensional vectors [2].
The Harris corner detector maybe one of the most vastly utilized detector [3], designed in 1988.It is depend on the eigenvalues of the second moment matrix. Although, Harris corners are not scale invariant. This method was refined by Mikolajczyk and Schmid [4] by creating sturdy and scale-invariant feature detectors with high repeatability, that they produce Harris Laplace and Hessian-Laplace. They choose the location by using a (scale adapted) Harris measure or the Hessian matrix determinant, while use Laplacian to choose the scale.

In 1999 D. G. Lowe [5] proposed a way to detect distinct invariant features from images which is called SIFT, later it uses to possible matching between various viewpoints of same object or scene, yet calculation time of SIFT made much slower. Henceforth researchers need to build up a detector and descriptor which carries on well harmony between speed and exactness. As a lead to 2006 H. Bay, T. Tuytelaars, and L. Van Gool [1] projected the Speeded up Robust Feature approach that creates efficient computation in robust way.

In this survey paper one first explain the SURF, within this section one basically clarify the SURF detector, descriptor and descriptor matching steps. The detector is relies on the Hessian matrix which provides better computation time and accuracy. SURF descriptor relies on the responses of Haar wavelet also it is computed by using integral images. As a way to detect interest points by the determinant of Hessian, first it is necessary to introduce the concept of a scale space [2]. And then expands the descriptor matching between different viewpoints of images.

2. Speeded up Robust Feature

SURF is an approach to detect and describe native features in image and finding the interest points in particular images that have totally different size of images and different viewpoints, completely different depths, scale changes and invariability during rotation, and strong to alternative typical geometric and mensuration transformation. It primarily uses for scale and rotation invariant feature transformation. The SURF approach consists of three major steps.

1. Detection: Interest point identification.
2. Description: Feature vector extraction that related to every interest points.
3. Descriptor matching: Verify correspondence between descriptors in two views.

The flowchart of SURF approach is described in the Figure1.

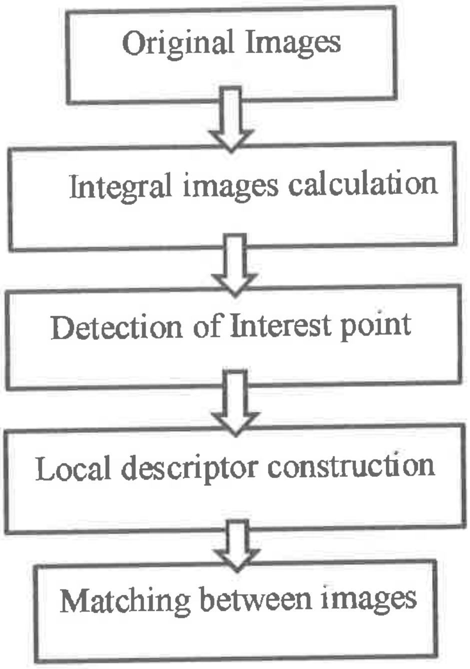


Figure1: SURF detector steps [6].

2.1 Interest Point Detection

In SURF, there are two steps associated with the detection process:

1. Integral images

One purpose that the SURF model is very fast to compute is the image space that all calculations are executed in. Rather than traditional RGB or grayscale images SURF use the integral image space to determine the interest points, Integral Image also known as summed area tables which is a way to represent the image. The main task of an integral image is computing the area of an upright rectangular region. It includes the total of severity values of each pixel in original image I inside the rectangular region created by origin O= (0, 0) and any point X=(x, y). It gives rapid calculation of box type convolution filters [7].

$$I\Sigma(X)\ \sum_{i=0}^{i\leq x}\sum_{j=0}^{j\leq y} I(i,j) \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

2. Fast Hessian Detector

The core concept of SURF is that the interest points are detected by the utmost determinant of the Hessian matrix. The Hessian matrix $\mathcal{H}(X, \sigma)$ in X for a point X=(x,y) of an image I, at particular scale can be seen as follows:

$$H(X, \sigma) = \begin{bmatrix} Lxx(X,\sigma) & Lxy(X,\sigma) \\ Lxy(X,\sigma) & Lyy(X,\sigma) \end{bmatrix} \qquad (2)$$

Where Lxx (X, σ) is the convolution of the Gaussian second order derivative with the image I in point X and similarly for Lxy(X, σ) and Lyy(X, σ)[8]. The SURF approximates the Hessian matrix determinant $HL(x, \sigma)$ by utilizing box filters, the 9 x 9 box filters shown in figure 2, are approximations of a Gaussian with $\sigma = 1.2$ and represent the lowest scale. They will indicated by Dxx, Dyy and Dxy. The grey regions are equal to zero.

$$\text{Det(Happrox)} = \text{Dxx Dyy} - (\text{wDxy})^2 \qquad (3)$$

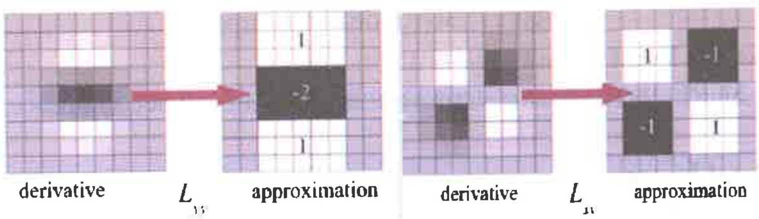


Figure 2: The Gaussian second order derivative and the approximation for the second order Gaussian derivative [1].

For the Hessian's determinant, the proportional weight w of the filter responses is employed for balancing the expression. Interest points have to be compelled to be discovered at totally various scales since the search of symmetries usually needs their comparison in images wherever they are seen at various scales.

SURF utilizes the integral image to approximate the various levels of scale space by adapting the box filters magnitude instead of the main image. The result of the 9 x 9 filter is taken into account as the primary scale layer, for that one it indicates as scale s = 1.2. Then, subsequent layers can be gained through image filtration by progressively larger masks, taking into consideration the distinct type of integral images, also the particular framework of the box filters. Eventually, interest points is found by non-maximum suppression in a very 3×3×3 neighborhood around every sample point. The response value of the characteristic point is larger than the response value of 26 adjacent points, then the position and scale information of the feature points is obtained [7].

2.2 Interest Point Description

The main objective of a descriptor is to produce a completely distinct and sturdy characterization of a feature, the creation of a descriptor depend totally on the region that surround the interest point. The SURF descriptor depends on Haar wavelet responses and it is computed expeditiously by using the integral images [5] including two stages: determine main direction and build descriptors.

2.2.1 Orientation Assignment

As a way for being constant to the rotation of images, the Haar wavelet responses can be calculated in both directions horizontal and vertical in a round neighborhood with six s radius surround the feature point, in which s is the scale where the feature point was detected. Haar wavelet describes the distribution of the intensities of pixel in a scale dependent neighborhood around every interest point that detected by fast Hessian. These responses appeared as vectors. After that each vector of both directions x and y of the Haar wavelet responses are summed inside a window that has a coverage angle of size 60 degree surround the feature point, figure 3 represent this filter the dark parts have the weight -1 and the light parts +1.The longest vector is the dominant orientation of the feature point.

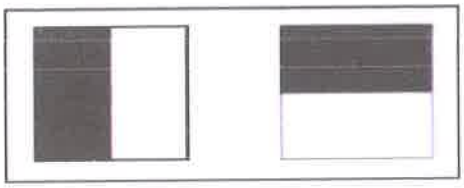


Figure 3: The filters of Haar wavelet responses in the two directions horizontal (left) and vertical (right) [1].

2.2.2 Building SURF descriptor with Haar Wavelet Responses

To extract the descriptor, first build a square area of 20s size and then the interest area is divided into 4×4 square sub-areas with 5×5 consistently spaced sample points inside and then the Haar wavelet response in x-direction dx (The Haar wavelet response in horizontal direction) and the Haar wavelet response in y direction dy (The Haar wavelet response in vertical direction) is computed as shown in figure 4. First the dx and dy responses are weighted by the use of Gaussian kernal placed in the center of the interest point. After that , the dx and dy responses are collected in every sub-area so that the initial set of inputs in the feature vector are formed .So as to get information concerning the polarity of the intensity changes, the total of absolute value of the responses is extracted. Thus, each sub-area is created a 4-dimensional vector,

$$V = (\sum d, \sum dy\ , \sum|dx|\ ,\sum|dx|)\ldots\ldots\ldots\ldots\ldots\ldots\ (4)$$

Connect the descriptor vector for every 4x4 square sub-areas, the feature vector dimension is 64. An oriented quadratic grid with 4 x 4 square sub-regions is set over the interest point (left) in order to build the descriptor. For every square, the wavelet responses are calculated. The 2 x 2 sub-divisions of every square relate to the specific fields of the descriptor. These are the totals dx, |dx|, dy and |dy| calculated comparatively to the orientation of the grid.

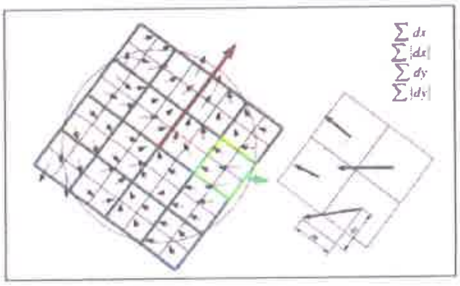


Figure 4: Building the descriptor [1].

2.3 Feature point matching

Descriptor matching procedure fulfills by means of SURF descriptor vectors that are applied to match between distinct viewpoints of images. This stage is most often relies on a distance between the vectors like Euclidean or Mahalanobis distances. SURF calculation time fundamentally depend on the dimension of the descriptor vectors. Therefore, lower numbers of dimension vectors are more suitable for its speed [2].

In the space of integral image, the trace of a Hessian matrix, or the Laplace, gives important information on the blob-to-background relation in the following neighborhood. To be more specific, the sign of the Laplace is an indicator of the brightness between the blob and the background. A positive Laplace shows a bright blob on a dark background, while a negative Laplace shows a dark blob on a bright background [1].

3. SURF Properties

SURF has different properties these are:

1. A very fast interest point detector and descriptor.
    - Keeping the performance similar with other detectors.
    - High repeatability (reliability of discovering same interest points under different viewing status).
2. Based on principles applied in David Lowe's SIFT, however the performance is higher.
3. Tested in real-world applications.
4. Take advantage of the integral image.
5. The SURF detector is an approximation to the Hessian.
6. Reuse the computations needed for detection in descriptor calculation.
7. Preserve robustness to rotation and scale illumination change.
8. About 2x faster than DoG 10x faster Hessian-Laplace detector.

4. SURF Advantages and Drawbacks

SURF has several advantages that include:

- Dealing with serious blurring.
- Dealing with image rotation.

- The major usefulness of SURF is the faster extraction of interest points, because of the usage of integral images.

While SURF drawbacks are:

- Dealing with viewpoint change.
- Dealing with illumination change.
- Not stable to rotation.

The characteristics and drawbacks of SIFT, PCA-SIFT and SURF are given within table1 [9].

METHOD	TIMES FOR 10 MATCHES	SCALE (NUMBER OF MATCHES UNDER SCALE)	ILLUMINATION (THE AVARAGE OF REPEATABILITY)	AFFINE (THE AVARAGE OF REPEATABILITY)
SIFT	2.14806e+007 (ms)	303 matches	21%	47%
PCA-SIFT	2.09696e+007 (ms)	85 matches	25%	15%
SURF	3304.97(ms)	418 matches	31%	54%

5. Comparison Between SIFT and SURF

SIFT and SURF are two local feature matching technique that ended up being most promising because of high achievement also have been utilized in several applications.

i. Within SIFT method, first the keypoints that are constant at illumination, scale, and rotation will be determined, thereafter feature vectors that correspond to every keypoints is calculated. As indicated by H. Bay [1] SURF method works in a way identical to SIFT however SURF follows another way in the sake of processing in the whole stages. Also H. Bay [1] research clears that SURF is considered as an improved method of SIFT.

ii. Within SURF method, there are 3 major processes these are detection of interest point, description of local neighborhood and matching. SURF utilizes a detector that depend on the Hessian matrix to locate the interest point. The Hessian matrix determinant is utilized as a degree of domestic alteration around the point, points are selected wherever this determinant is most extreme. Then the various categories of keypoints are classified by nearest neighbor approach [10]. Table (2) shows a comparison between the characteristic of SIFT and SURF.

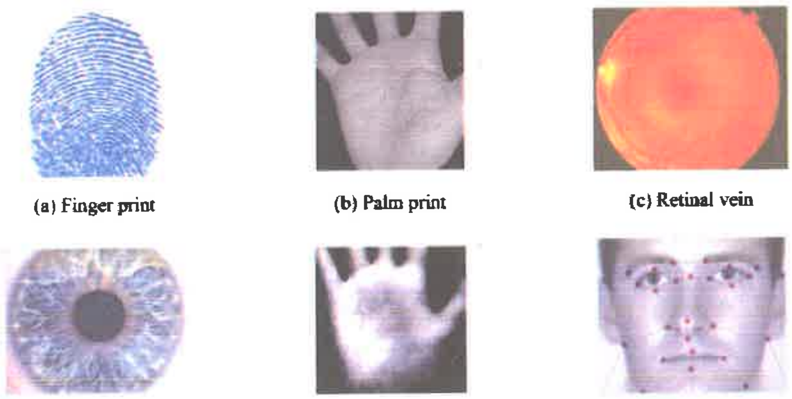
Table2: The characteristics of SIFT and SURF [10].

	SIFT	SURF
Keypoint Detection	Diverse scale image convoluted with Gaussian capacity	Original Image is convoluted with Different scale box Filter
	Nominee keypoint detect extrema in Difference of Gaussian space	Candidate keypoints are determined using Hessian matrix
Keypoint Description	Gradient capacity of a square region is computed with extreme gradient power as the major direction	A Haar wavelet response is utilized to compute all sectors in round region
	The Extraction of feature is done by partitioning a 16×16 area to 4×4 subarea, for every subarea a gradient Histogram is made.	The Extraction of feature is done by dividing a 20×20 area into 4×4 subarea and a Haar wavelet Response is computed.
Sizes	128 element	64 element

6. Application of SURF on Retinal image Systems

The retina is a tissue composed of several layers which covers the inside of the eye and converts the entering light into a neural signal that undergoes for more processing in the visual cortex of the brain.

Different biometric frameworks have been made depending on fingerprint, palm print, hand geometry, veins, face, iris, retina and ear figure 5 give some examples. In the recent years Retinal recognition has gain an increasing attention because of its ability to solve security problems as a result of its comprehensiveness, distinction, time-invariance and that it cannot be faked. As retinal patterns have very particular features, these features that taken from retina can recognize persons efficiently, even between genetically identical twins. The pattern will remain invariant during the person's life, except when a dangerous disease occurs within the eye. Most popular sickness like diabetes don't alter the pattern in a way that its structure can be influenced. Therefore, retina in general is secure and useful biometric feature for confirming persons [11].



(a) Finger print (b) Palm print (c) Retinal vein

(d) Iris print (e) Palm vein (f) Face recognition

Figure 5: Some biometric patterns for identification and verification.

6.1 Verification System Based on Retina and SURF

Takwa Chihaoui, Hejer Jlassi, Rostom Kachouri, Kamel Hamrouni, and Mohamed Akil [12] proposed a new recognition framework based on retina and SURF. Within the suggested framework, they don't use segmentation techniques that raise computational time. So as to enhance the quality and decrease the time of processing, a preprocessing stage depend on Optical Disc interest Ring (ODR) technique which extracts an interest ring that surround the optical disc is required. Then, the speedy extraction of SURF interest points can be applied to supply a characterization of interest points of the image of retina.

In final stage within the process of verification, the matching among the interest points of the reference image and the interest points of the input image is done by using the test g2NN.

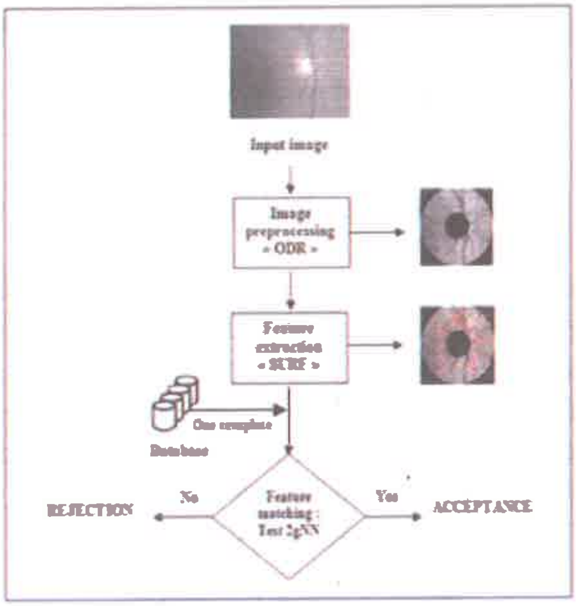
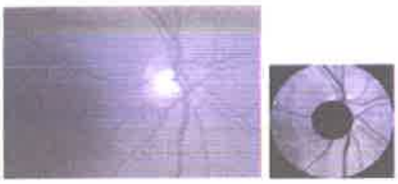


Figure 6: Flowchart of the proposed retinal system.

6.1.1 System Overview

Interest ring extraction: so as to overcome the issues of retinal images, they use recently proposed method known as ODR [13] in order to preprocess the retinal image for enhancing quality of the image and take out an interest ring that focused on the middle of the optical disc which is considered an interest area of this image.

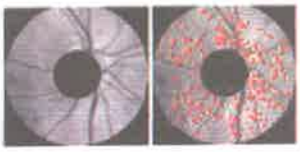


(a) (b)

Figure 7: The preprocessing stage: (a) Input image of retina

(b) The Optical Disc interest Ring output.

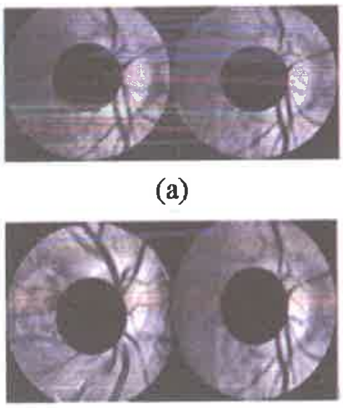
Feature extraction: SURF approach [1] is applied to give a characterization of the retinal image features. The result of feature extraction stage is a feature vector of 64 elements for every detected point. Figure 8 shows the distribution of SURF interest points on the retinal image that have been preprocessed.



(a) (b)

Figure 8: The distribution of SURF interest points: (a) The extraction of interest ring (b) The characterization of SURF interest points.

The strategy of feature matching: In the SURF space the step of matching the feature vectors of each interest point is performed so that the similar regions in the sampled image can be identified. So, they use a matching technique referred to as g2NN check to discover the better nominee that depend on the distance with the first and the second most comparable interest point.



(a)

(b)

Figure9: The g2NN test authentication cases
(a) The clients (licensed) case (b) The attacks (unlicensed) case

6.2 Retinal Image Registration method Based on SURF

Zhitao Xiao, Wan Zhu, Fang Zhang, Jun Wu, Lei Geng and Wen Wang [14] proposed a registration method based on speed-up robust feature, which is based on the original retinal image instead of vessels. The algorithm mainly consist of three parts: Firstly, preprocessing is performed on the retinal images, Secondly, from the reference image and the floating image SURF feature points are extracted, and the descriptors are generated later. Finally, the spatial transformation of the floating image is achieved.

6.2.1 System Overview

Preprocessing: The three channels of color retinal images are extracted and the green channel of the retinal image is selected for subsequent processing. Then a contrast Enhanced Adaptive Histogram Equalization (CLAHE) [15] technique is utilized to enhance the contrast of the green channel of the retinal image, the result is shown as Figure10.

Figure10: CLAHE image.

Features matching: Extract feature points from retinal image using detector of SURF is relies on integral images, Hessian matrix, and scale space theory. After the establishment of feature descriptor, they used Euclidean distance to recognize the initial matching of feature points, then matching points are filtered using the random sample consensus algorithm. First they Calculate the Euclidean distance of the feature points on the floating image to the reference image, then the distance sets are obtained, if the minimum Euclidean distance ratio and the second smallest Euclidean distance ratio is lower than the threshold, they think that the feature point and the minimum Euclidean distance of feature points are matched.

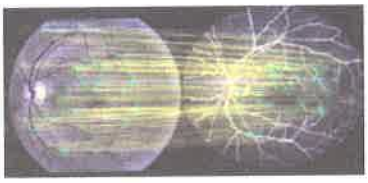


Figure11: Matching results.

Transformation model and parameter estimation: There are three types of models of retinal image: similar, affine and quadratic curve model, the performance of the affine model have the higher percentage. Compared with the quadratic curve model, affine transformation model won't lose an excessive amount of information so they choose this model to carry on the spatial transformation and the registration results are shown as Figure 12.

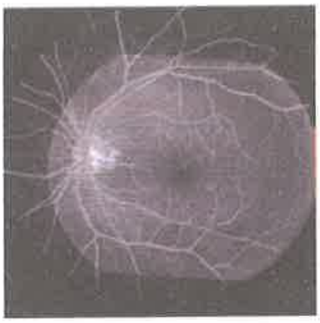


Figure12: Registration results.

7. Conclusions

The SURF technique is mainly used for detecting and describing interest points of the image in robust way by utilizing the integral images, fast Hessian detector, and the sum of wavelet responses for the orientation assignment to make better performance than other similar methods like SIFT, PCA SIFT. SURF is three times faster and has a higher improvement in recall precision than SIFT technique, however it is very poor at handling viewpoints of images but can be computed and compared extremely faster.

Speeded up Robust Feature is used in several applications of computer vision such as verification, registration and recognition systems like biometrics to make higher efficiency in detecting and describing the sturdy and special features that can be utilized to recognize a person or check the identity of a person.

References

[1] H. Bay, T. Tuytelaars, and L. V. Gool (2006), “SURF: Speeded Up Robust Features”, Computer Vision–ECCV.
[2] W.O.K.A.S.Wijesinghe (2010)," speeded up robust feature in computer vision systems", Reg No: CS/036, Index No: 10000364.
[3] C. Harris, M. Stephens (1988), "A combined corner and edge detector", in: Proceedings of the Alvey Vision Conference, pp. 147–151.
[4] K. Mikolajczyk, C. Schmid (2001), "Indexing based on scale invariant interest points", in: ICCV, vol. 1, pp. 525–531.
[5] D. G. Lowe (1999), “Object recognition from local scale-invariant features,” Proceedings of the Seventh IEEE International Conference on Computer Vision, pp. 1150–1157 vol.2.
[6] Vimal Singh Bind, Priya Ranjan Mudul and Umesh Chandra Pati (2013), "A Robust Technique for Feature-based Image Mosaicing using Image Fusion‖", International Journal of Advanced Computer Research (IJACR), Vol.3, Issue-8, pp.263-268, India.
[7] Utsav Shah,Darshana Mistry,Yatin Pate(2014), " Survey of Feature Points Detection and Matching using SURF, SIFT and PCA-SIFT", Journal of Emerging Technologies and Innovative Research (JETIR).
[8] Jacob Toft Pedersen (2011)," Study group SURF: Feature detection & description", SURF: FEATURE DETECTION & DESCRIPTION".
[9] Juan L., Gwun O. (2010), " A comparison of sift, pca-sift and surf ", International Journal of Image Processing (IJIP).
[10] Sheena S, Sheena Mathew (2016)," A comparison of SIFT and SURF algorithm for the recognition of an efficient iris biometric system", International Journal of Advanced Research in Computer and Communication Engineering.
[11] Amin Dehghani, Zeinab Ghassabi, Hamid Abrishami Moghddam and Mohammad Shahram Moin (2013)," Human recognition based on retinal images and using new similarity function", EURASIP Journal on Image and Video Processing.
[12] Takwa Chihaoui, Hejer Jlassi, Rostom Kachouri, Kamel Hamrouni, Mohamed Akil (MAR 2016), "Personal verification system based on retina and SURF descriptors", 13th IEEE International MultiConference on Systems, Signals & Devices, Leipzig, Germany.

[13] eyed Mehdi Lajevardi, Arathi Arakala, Stephen A Davis, and Kathy J Horadam(09/2013), "Retina Verification System Based on Biometric Graph Matching", IEEE Transactions on Image Processing; 22(9):3625-3635.

[14] Zhitao Xiao, Wan Zhu, Fang Zhang, Jun Wu, Lei Geng, Wen Wang (2016),"Multimodal Retinal Image Registration method Based on Speed-up Robust Feature", 2nd Workshop on Advanced Research and Technology in Industry Applications.

[15] Setiawan AW, Mengko T R, Santoso O S, et al (2013),"Color retinal image enhancement using CLAHE". International Conference on ICT for Smart-Society (ICISS). Jakarta, pp.215-217.