



FRONTIERS IN ELECTRONICS AND COMMUNICATION ENGINEERING

ISSN: (3065- 4106)



editor.fece@gmail.com



<https://multisciajournals.com/journals/index.php/fece>

A SIFT and SURF Comparison

Elena I. Kostyukova, Youngsighn
Department of ECE

Article Info

Received: 28-02-2025 Revised: 08-03-2025 Accepted: 16-03-2025 Published: 27-03-2025

Abstract: In many applications, accurate, reliable, and automatic image registration is a crucial task. Feature detection, feature matching, derivation of transformation function based on related features in images, and reconstruction of images based on derived transformation function are necessary phases in image registration/alignment. Accurate feature detection and matching are necessary for a registered image to be accurate. Therefore, in many picture applications, such as image registration, computer vision, image mosaic, etc., these two intermediate processes are crucial. Scale Invariant Feature Transform (SIFT) and Speed Up Robust Features (SURF) are two distinct approaches for scale and rotation invariant interest point/feature detector and descriptor that are presented in this study. Additionally, it offers a method for obtaining unique invariant features from pictures that can be utilized to reliably match various perspectives of a scene or item.

Keywords: SIFT, SURF, NCC, feature matching, and feature detection.

INTRODUCTION

- In order to extract unique invariant features from images that can be invariant to image rotation and scaling, Lowe (2004) introduced SIFT. After that, image mosaic, recognition, retrieval, and other applications made extensive use of it [3]. Using integral images for image convolutions and the Fast-Hessian detector, Bay and Tuytelaars (2006) accelerated robust features. According to their experiments, it functions well and is speedier.
- Many computer vision applications include picture matching tasks that involve determining the correspondences between two images of the same scene or item. Object recognition, camera calibration, and image registration are just a handful. There are two primary stages to this paper's description of distinguishing characteristics from photos. Initially, "key points" are taken from identifiable spots in the pictures, like corners, edges, and blobs. High repeatability is a must for key point detectors. After that, neighborhood regions are selected around each important point, and each region's unique feature descriptors are calculated.
- Features that can reliably match images from various perspectives are extracted from photos for image matching. The procedure involves the extraction and storage of feature descriptors from sample photos. This descriptor must be unique while also being resistant to noise and detection errors. Lastly, the feature descriptors are compared across several photos. Euclidean distances can be used as the basis for feature descriptor matching. The techniques overview is covered in Section 2 of this study, the experimental results are shown in Section 3, and the paper's conclusions are presented in Section 4.

I. OVERVIEW OF METHODS

SIFT ALGORITHM OVERVIEW

In 2004 [6], Lowe presented the SIFT (Scale Invariant Feature Transform) technique, which has strong robustness and can tackle picture rotation, scaling, affine deformation, viewpoint shifts, noise, and lighting changes.

There are four primary steps in the SIFT algorithm: Detection of Scale Space Extrema, Localization of Key Points, Assignment of Orientation, and Generation of Descriptions. Using scale space extrema in the DoG (Difference-of-Gaussian) functions with varying values of σ , the first step is to locate and scale critical points. The DoG function is convolved of images in scale space separated by a constant factor k , as shown in the equation below.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y)$$

where I is the image and G is the Gaussian function. The DoG is now obtained by subtracting the Gaussian images, and then the Gaussian images subsample by factor 2 to obtain the DoG for the sampled image. To find the local maxima and minima of $D(x, y, \sigma)$, a pixel comparison of a 3x3 neighborhood is performed.

Key point candidates are localized and refined in the key point localization step by removing the key points that rejected the low contrast points. The orientation of the key point is determined using the local image gradient in the orientation assignment step. The local image descriptor for each key point is calculated during the description generation stage using the image gradient magnitude and orientation at each image sample point in a region centered at the key point [2];

Each sample has eight orientation bins and a 4x4 array location grid, creating a three-dimensional histogram of gradient location and orientation. That is the key point descriptor's 128-element dimension.

Construction Of SIFT Descriptor

The key point descriptor computation is shown in Figure 1. To choose the degree of Gaussian blur for the image, the size of the key point is used to sample the image gradient magnitudes and orientations around the key point location [6]. The gradient orientations are then rotated with respect to the key point orientation in order to establish orientation invariance. Small arrows are shown at each sample location on the left side of Figure 1.

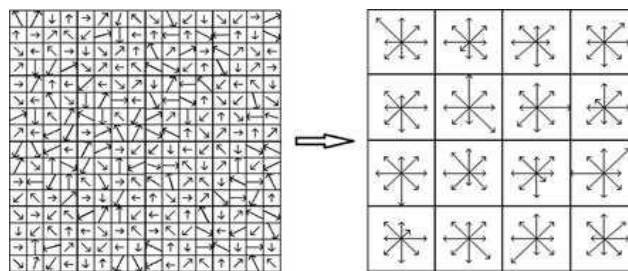


Figure 1: SIFT Descriptor Generation

Figure 1's right side displays the key point descriptor. By generating orientation histograms across 4x4 sample sections, it permits notable shifts in gradient positions. Each orientation histogram has eight directions in the picture [6], and the length of each arrow represents the magnitude of that histogram entry. In order to contribute to the same histogram on the right, a gradient sample on the left can move up to four sample places. Thus, each sample has eight orientation bins and a 4x4 array placement grid. That is the key point descriptor's 128-element dimension.

SURF Algorithm Overview

The feature detector is based on the Hessian matrix, and the SURF (Speed Up Robust Features) technique is based on multi-scale space theory. because the Hessian matrix performs well and is accurate. The supplied point in image I is $x = (x, y)$. The Hessian matrix $H(x, \sigma)$ in x at scale σ can be defined as

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\sigma}) = \begin{bmatrix} L_{xx}(\mathbf{x}, \boldsymbol{\sigma}) & L_{xy}(\mathbf{x}, \boldsymbol{\sigma}) \\ L_{yx}(\mathbf{x}, \boldsymbol{\sigma}) & L_{yy}(\mathbf{x}, \boldsymbol{\sigma}) \end{bmatrix} \dots\dots\dots (2)$$

where $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$ are the convolution results of the second order derivative of the Gaussian filter in point x , respectively.

Using the image I ∂_x 2 , ∂^2 $g(\sigma)$

For higher levels in the pyramid, SURF generates a "stack" without 2:1 down sampling, producing images with the same resolution. SURF uses a box filter approximation of second-order Gaussian partial derivatives to filter the stack because it uses integral pictures [3]. Rectangular box filters can be computed in almost constant time thanks to integral pictures. The Gaussian second order partial derivatives in both the y and xy directions are displayed in Figure 2.

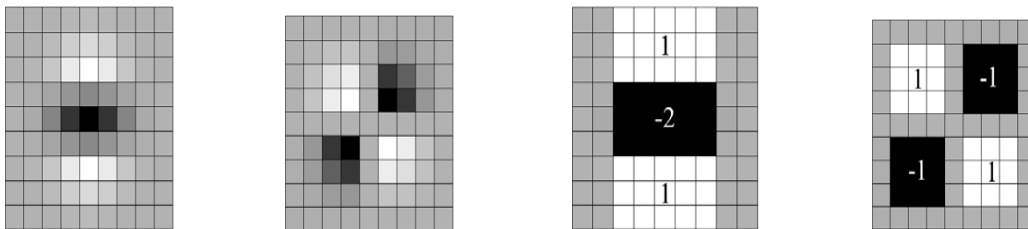


Figure 2. The Gaussian second orders partial derivatives in y-direction and xy-direction [4].

When compared to other descriptors, SIFT performs well. Similar characteristics serve as the foundation for the suggested SURF description. Fixing a repeatable orientation using data from a circular area surrounding the point of interest is the first stage. The SURF descriptor is then extracted from a square region that has been constructed and aligned with the chosen orientation. It computes the Haar-wavelet responses in the x and y directions, as seen in figure 3, to be invariant to rotation.



Figure3 Haar wavelet types used for SURF

II. EXPERIMENTAL RESULTS

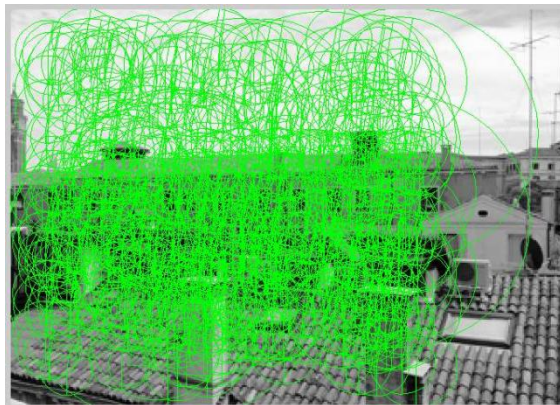
III. Two photographs are used as the experimental data to confirm the algorithm's efficacy, as seen in figure 1 (a) image1: 640×478, 153 KB, and (b) image2: 640×478, 127 KB. The Intel Core i-3 3210, 2.3 GHz processor, 4 GB RAM, and Windows 7 operating system are used for the research. The SIFT and SURF algorithms are used to identify features in both images. The SIFT-detected features in images 1 and 2 are displayed in Figure 4(c) and (d), respectively. 892 features are found in picture 1, whereas 934 characteristics are found in image 2. The SURF algorithm-detected features from the original images 1 and 2 are displayed in Figures 4 (f) and (g), respectively. It can be seen that image 1 has 281 features while image 2 has 245 features.



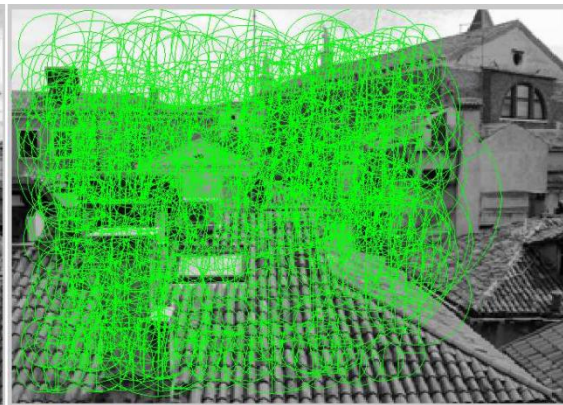
(a) Original image1



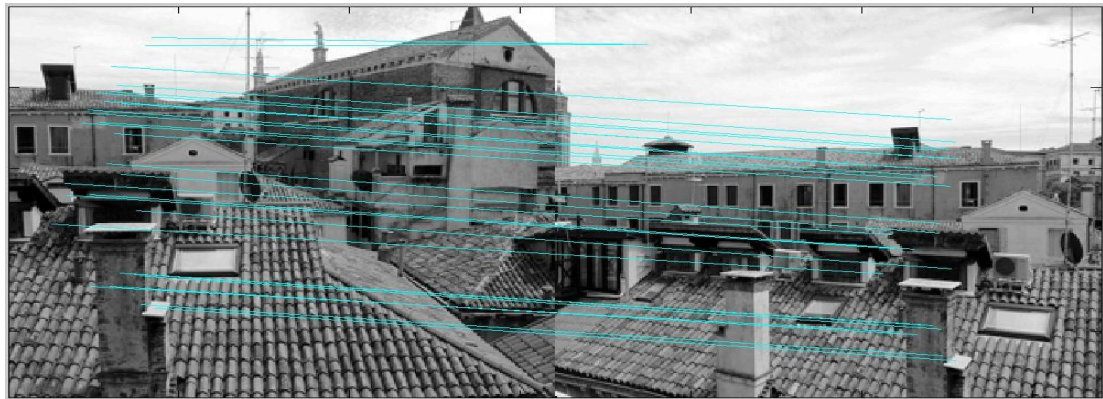
(b) Original image2



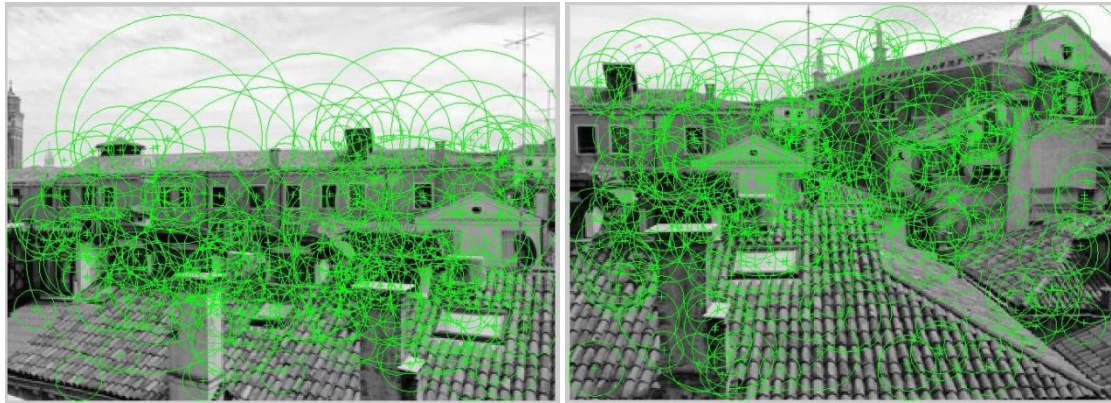
(c) Detected features in image1 using SIFT



(d) Detected features in image2 using SIFT

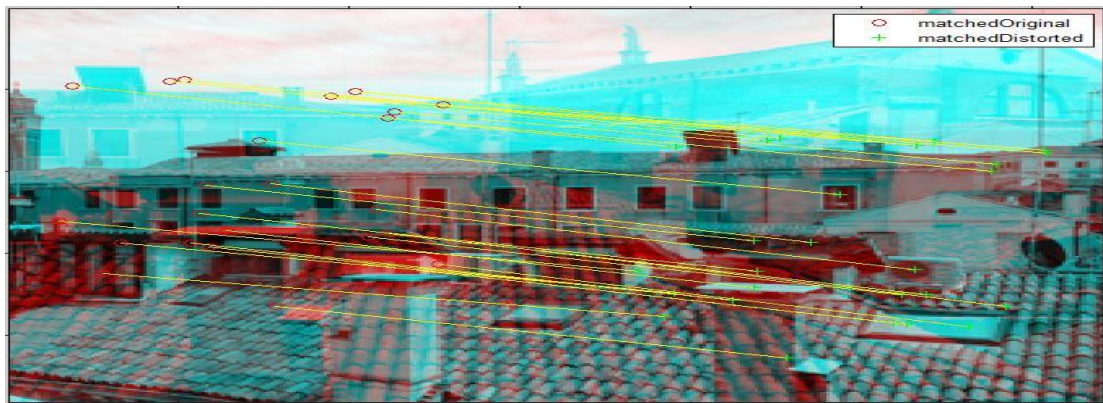


(e) Matching pairs identified between the image1 and image2



(f) Detected features using SURF in image1

(g) Detected features using SURF in image2



(h) Matching pairs identified between the image1 and image2

Figure 4(e) displays 41 matched features, while Figure 4(h) displays 28 matched points. Here, feature matching is done using the Normalized Cross Correlation approach. Table 1 summarizes the experimental findings.

TABLE 1: Comparisons of results of SIFT and SURF algorithm

Algorithm	Detected feature Points		Matching feature point	Feature matching Time
	Image1	Image2		
SIFT (Scale Invariant Feature Transform)	892	934	41	1.543 s
SURF (Speed Up Robust Feature)	281	245	28	0.546 s

IV. CONCLUSIONS

V. Two feature detection techniques for image registration have been assessed in this paper. According to the experimental findings, SIFT has identified more features than SURF, although it suffers with speed. Similar to SIFT, the SURF is quick and performs well. Our goal for the future is to implement these algorithms for video registration.

ACKNOWLEDGMENT

The authors of this study are grateful to Bay and Tuytelaars for their assistance in creating SURF code and their valuable advice on this subject, as well as to David Lowe for his assistance in creating SIFT code.

REFERENCES

- [1] Nabeel Younus Khan, Brendan McCane, and Geoff Wyvill, "SIFT and SURF Performance Evaluation against Various Image Deformations on Benchmark Dataset," International Conference on Digital Image Computing: Techniques and Applications, pp. 501-506, 2011.
- [2] "Automatic Image Registration using SIFT-NCC," Special Issue of the International Journal of Computer Applications (0975-8887), pp. 29-32, June 2012, Vini Vidyadharan and Subu Surendran.
- [3] "A Comparison of SIFT, PCA-SIFT, and SURF," Luo Juan and Oubong Gwun, International Journal of Image Processing (IJIP), Vol. 3, Issue 4, pp. 143-152.
- [4] Herbert Bay, Luc Van Gool, and Tinne Tuytelaars, "SURF: Speeded Up Robust Features," pp. 1-14.
- [5] Yong-Ho Moon and Seok-Wun Ha, "Multiple Object Tracking Using SIFT Features and Location Matching," International Journal of Smart Home, Vol. 5, No. 4, October 2011, pp. 17-26.
- [6] "Distinctive Image Features from Scale-Invariant Keypoints," by D. Lowe, was accepted for publication in the International Journal of Computer Vision, page 1-28, 2004.
- [7] Hongbo Li, Ming Qi, and Yu Wu, Journal of Theoretical and Applied Information Technology, Vol. 42, No.2, pp. 281-286, August 2012, "A Real-Time Registration Method Of Augmented Reality Based On Surf And Optical Flow."