Computers and Electrical Engineering 000 (2017) 1–13

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

# Feature points for multisensor images☆

Sajid Saleem[a,*], Abdul Bais[b], Robert Sablatnig[c], Ayaz Ahmad[d], Noman Naseer[e]

[a] Faculty of Engineering and Computer Sciences, National University of Modern Languages, Islamabad, Pakistan
[b] Faculty of Engineering and Applied Science, University of Regina, Canada
[c] Computer Vision Lab, Institute of Computer Aided Automation, Vienna University of Technology, 1040 Vienna, Austria
[d] Department of Electrical Engineering, COMSATS Institute of Information Technology, Wah Cantt, Pakistan
[e] Department of Mechatronics Engineering, Air University, Islamabad, Pakistan

ARTICLE INFO

*Article history:*
Received 4 October 2016
Revised 12 April 2017
Accepted 14 April 2017
Available online xxx

*Keywords:*
Feature point detectors
Feature point descriptors
Multisensor images
Image matching

ABSTRACT

Feature points are effective for wide range of computer vision applications. In the last two decades, a large number of feature point detection and description algorithms have been proposed. All these algorithms are implemented differently but have a common objective to detect and describe feature points invariant to scale, rotation, intensity, and affine variations. Several comparative studies of feature points have been reported in literature. These studies are either application specific or deal with common type of transformations and deformations. Additionally, they primarily focus on evaluation of feature points on gray scale or RGB images. In contrast, this paper presents a comparison of feature points on multisensor images, which possess non linear intensity changes. The objective is to identify robust feature points for image to image matching tasks on multisensor images. Six well known feature point detector and seventeen popular descriptor algorithms are compared. Experimental results obtained on four different image datasets show that the combination of Harris detector with BRIEF descriptor outperforms all other detector-descriptor combinations on multisensor images.



## 1. Introduction

A fundamental problem in computer vision is the detection, description, and matching of feature points [1], which forms the basis of many applications such as image registration, object recognition, image retrieval, 3D reconstruction, and object tracking. The detection of feature points deals with the extraction of corners, junctions, line intersections, blob like structures, etc from images [2]. Whereas the feature point description is the assignment of distinct descriptor vectors to feature points so that feature point correspondences can be established reliably between the images by matching the feature point descriptor vectors [3].

Gray scale and RGB images are extensively used in wide range of computer vision applications [4]. Therefore, existing feature point detector and descriptor algorithms are designed for these image formats. These algorithms aim to extract feature points in such a manner that common types of transformations and deformations (such as rotation, scale, affine, projective, and illumination changes) can be overcome efficiently [5].

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. G. Botella.

* Corresponding author.
*E-mail address:* sajidslm@yahoo.com (S. Saleem).

http://dx.doi.org/10.1016/j.compeleceng.2017.04.032
0045-7906/© 2017 Elsevier Ltd. All rights reserved.

Recently, multispectral and multisensor images have received a lot of attention [6,7]. Such images provide more distinct visual information and higher spectral resolution than typical gray scale and RGB images. Cross spectral stereo matching [8], cross spectral image registration [9], cross spectral face recognition [10], and remote sensing [11] are a few applications, where multisensor and multispectral images are used. Feature points are used in these applications to fuse, register, or match different spectral band images.

The demand for robust feature points to deal with non linear intensity changes between spectral band images, has resulted in numerous studies, where existing feature points are either modified [9,11,12] or new feature points are proposed [8]. Despite this, the suitability of feature points is still a problem. Because such feature points are not suitable for each and every application that involves multisensor and multispectral images. A comprehensive comparison of feature points is required. In this regard, this paper presents a comparison to identify robust feature points for multisensor images. Our comparison is different from other studies, which either focus on common types of image transformations and deformations [2,3,5,13] or provide limited experimental results on multisensor images [8,12,14].

The rest of this paper is organized as follows: Section 2 gives an overview of classical and state of the art feature points. Section 3 presents evaluation criteria and introduces four different image datasets. Section 4, presents a comparison of feature point detectors and descriptors. Finally, the paper is concluded in Section 5.

## 2. Overview of feature point detectors and descriptors

Effective and efficient detection and description of feature points is a well studied problem in computer vision. The detection of feature points can be traced back to [15], where a local image window is slided over the whole image to detect corners as feature points. With each slide the window is shifted in various directions to determine changes in the average pixel intensity. Several modifications to this method have been proposed [2] and also various algorithms have been suggested to construct distinct descriptors for detected corners [5]. This section gives an overview of six and seventeen well known feature points detector and descriptor algorithms, respectively.

### *2.1. Harris*

The corner detector of [15] is modified in [16]. Instead of using a local window, a smoothed circular window is used to deal with noise and other image deformations. This modified algorithm is known as Harris corner detector. Several modifications to this algorithm have been proposed [2].

### *2.2. SIFT*

Scale Invariant Feature Transform (SIFT) [1] uses Difference of Gaussian (DoG) function and a Hessian matrix for feature point detection. The input image is convolved with Gaussian filters and then down sampled. An image pyramid is obtained. The adjacent scale images in the image pyramid are subtracted to realize the DoG implementation. Each pixel of the subtracted images is analyzed in the current and adjacent scales for local extrema detection as potential feature points. Such points are scale and rotation invariant. A 3D quadratic function is fitted to each detected extrema to accurately localize extrema both in spatial and scale domains. Finally, the Hessian matrix is applied to eliminate extrema lying on edges, because the DoG function gives strong responses along the edges.

SIFT descriptor is based on image gradients around SIFT feature points. The gradients are spatially divided into $4 \times 4$ location bins. For each location bin, a histogram of oriented gradients is computed. The histograms are concatenated over all location bins to obtain a SIFT descriptor of size 128. The SIFT descriptor is widely used. Several modifications of the SIFT descriptor have been proposed, such as Gradient Location and Orientation Histogram (GLOH) [3] and Color-SIFT [4].

### *2.3. GLOH*

GLOH is an extended version of SIFT descriptor. To obtain GLOH, image regions around feature points are convolved with directional gradients similar to SIFT. The image gradients are spatially divided into location bins using a log polar location binning scheme. A gradient histogram of 16 orientation bins is computed for each location bin. The histograms are concatenated to obtain a GLOH descriptor of size $16 \times 17 = 272$, where 17 represents number of location bins. The size of GLOH descriptor is reduced to 128 with Principle Component Analysis.

### *2.4. SURF*

Speeded Up Robust Features (SURF) [17] are based on integral images and the Hessian matrix. Integral images speed up the image convolution process and reduce the computation time whereas Hessian matrix identifies potential SURF feature points. The Hessian matrix is realized with the convolution of Gaussian second order derivative with the image. The determinant of Hessian matrix is computed at each pixel location and the locations where the determinant values are high are identified as SURF feature points.