Comparative Study of Different Image Feature Extraction Algorithm and Representation Techniques

PrashantAglave

Department Of Electronics Engineering,
Bharatratna Indira Gandhi College of Engineering,
Solapur Maharashtra

Vijaykumar S. Kolkure

Department Of Electronics Engineering,
Bharatratna Indira Gandhi College of Engineering,
Solapur ,Maharashtra

Abstract-Feature extraction is one of the most important step for image processing.The main objective of a feature extraction technique is to accurately retrieve features from the image*.*This paper compares three robust feature detection methods, they are, Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF) andOriented FAST and Rotated BRIEF (ORB).This paper gives an overall idea of general methods of Feature extraction and gives significance of ORB over SIFT and SURF algorithm. Use of ORB provides rotation invariance of both train image and query image.

Keyword

SIFT, SURF, ORB.

1. INTRODUCTION

Feature-based image matching is an important aspect in many computer based applications, such as object recognition, images stitching, structure-from-motion and 3D stereo reconstruction [2]. These applications require often real-time performance. Feature-based algorithms are well-suited for such operations. Different algorithms are used for image processing like Scale-invariant feature transform (SIFT), Speeded Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB). ORB is a very fast binary descriptor based on BRIEF, which is rotation invariant and resistant to noise.

ORB also uses learning method for de-correlating BRIEF features under rotational invariance, leading to better performance in nearest-neighbor applications. The recognize method for object recognition is Scale invariant feature transform (SIFT), which is popular for its invariance to scaling, rotation and illumination, is computationally complex due to its heavy workload required in local feature extraction and matching operation. Thus computer vision kind of applications demands high performance and low complexity solution and ORB provides better solution to it.

Feature detection is a low-level image processing operation. That is, it is usually performed as the first operation on an image, and examines every pixel to see if there is a feature present at that pixel. If this is part of a larger algorithm, then the algorithm will typically only examine the image in the region of the features. Many computer vision algorithms use feature detection as the initial step, so as a result, a very large number of feature detectors have been developed.

2.TYPES OF IMAGE FEATURES

A.Edges

Edges are points where there is a boundary (or an edge) between two image regions. In general, an edge can be of almost arbitrary shape, and may include junctions. In practice, edges are usually defined as sets of points in the image which have a strong gradient magnitude.

B. Corners / Interest points

The terms corners and interest points are used somewhat interchangeably and refer to point- like features in an image, which have a local two dimensional structure. The name “Corner" arose since early algorithms first performed edge detection, and then analyzed the edges to find rapid changes in direction (corners).

C. Blobs / regions of interest or interest points

Blobs provide a complementary description of image structures in terms of regions, asopposed to corners that are more point-like. Nevertheless, blob descriptors often contain a preferred point (a local maximum of an operator response or a center of gravity) which means that many blob detectors may also be regarded as interest point operators. Blob detectors can detect areas in an image which are too smooth to be detected by a corner detector.

D. Ridges

For elongated objects, the notion of ridges is a natural tool. A ridge descriptor computed from a grey-level image can be seen as a generalization of a medial axis. From a practical viewpoint, a ridge can be thought of as a one-dimensional curve that represents an axis of symmetry, and in addition has an attribute of local ridge width associated with each ridge point.

Base on the above types of features, there are different types of feature detection algorithms which is explained below.

3. FEATURE DETECTION

A. Sobel operator

The Sobel operator, sometimes called Sobel Filter, is used in image processing and computer vision, particularly within edge detection algorithms, and creates an image which emphasizes edges and transitions. it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. On the other hand, the gradient approximation that it produces is relatively crude, in particular for high frequency variations in the image.

B. Canny edge detector

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images.

C. Features from accelerated segment test

Features from accelerated segment test (FAST) is a corner detection method, which could be used to extract feature points and later used to track and map objects in many computer vision tasks. The most promising advantage of FAST corner detector is its computational efficiency. Referring to its name, it is fast and indeed it is faster than many other well-known feature extraction methods, such as difference of Gaussians (DoG) used by SIFT, and Harris.

4.FEATURE EXTRACTION

Once features have been detected, a local image patch around the feature can be extracted. This extraction may involve quite considerable amounts of image processing. The result is known as a feature descriptor or feature vector. In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction.

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

5.DIFFERENT FEATURE EXTRACTION METHODS

The different techniques used for feature extraction are:

1. Scale-Invariant Feature Transform (SIFT) [5]
2. Speeded Up Robust Features (SURF) [4]
3. Oriented FAST and Rotated BRIEF (ORB) [1] etc.

A. Scale-Invariant Feature Transform (SIFT):

SIFT (Scale Invariant Feature Transform) algorithm proposed by Lowe in 2004 [7] to solve the image rotation, scaling, and affine deformation, viewpoint change, noise, illumination changes, also has strong robustness. The SIFT algorithm has four main steps: (1) Scale Space Extrema Detection, (2) Key point Localization, (3) Orientation Assignment and (4) Description Generation.

The first stage is to identify location and scales of key points using scale space extrema in the DoG (Difference-of-Gaussian) functions with different values of σ, the DoG function is convolved of image in scale space separated by a constant factor k as in the following equation.

$$D(x,y,\sigma)=(G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y)=L(x,y,k\sigma) \quad (1)$$

Where, G is the Gaussian function and I is the image. Now the Gaussian images are subtracted to produce a DoG, after that the Gaussian image subsample by factor 2 and produce DoG for sampled image. A pixel compared of 3×3 neighborhood to detect the local maxima and minima of D(x, y, σ). Inthe key point localization step, key point candidates are localized and refined by eliminating the key points where they rejected the low contrast points. In the orientation assignment step, the orientation of key point is obtained based on local image gradient. In description generation stage is to compute the local image descriptor for each key point based on image gradient magnitude and orientation at each image sample point in a region centered at key point [6], these samples building 3D histogram of gradient location and orientation; with 4×4 array location grid and 8 orientation bins in each sample. That is 128-element dimension of key point descriptor.

Construction Of SIFT Descriptor:

Figure 5.1 illustrates the computation of the key point descriptor. First the image gradient magnitudes and orientations are sampled around the key point location, using the scale of the key point to select the level of Gaussian blur for the image [6]. In order to achieve orientation invariance, the coordinates of the descriptor, then the gradient orientations are rotated relative to the key point orientation. Figure 5.1 illustrated with small arrows at each sample location on the left side.

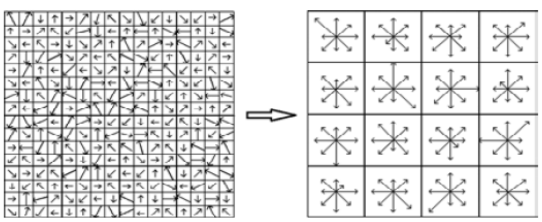

Figure 5.1 SIFT Descriptors

B.Speeded Up Robust Features (SURF)

SURF (Speeded Up Robust Features) [4] presents a novel scale- and rotation-invariant interest point detector and descriptor. It approximates or even outperforms previous schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster. This is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors (e.g., using a Hessian matrix-based measure for the detector, and a distribution-based descriptor); and by simplifying these methods to the essential. This leads to a combination of novel detection, description, and matching steps. The detector is based on the Hessian matrix [11, 1], but uses a very basic approximation, just as DoG is a very basic Laplacian-based detector. It relies on integral images to reduce the computation time hence called the 'Fast-Hessian' detector. The descriptor, on the other hand, describes a distribution of Haar-wavelet responses within the interest point neighborhood. Again, integral images are exploited for speed.

Figure 5.2 Haar wavelet types used for SURF

C .ORB

ORB is basically a fusion of FAST keypoint detector and BRIEF descriptor with many modifications to enhance the performance. ORB [1] builds on the well-known FAST keypoint detector and the recently-developed BRIEF descriptor; for this reason we call it ORB (Oriented FAST and Rotated BRIEF). Both these techniques are attractive because of their good performance and low cost. ORB includes,

- The addition of a fast and accurate orientation component to FAST.
- The efficient computation of oriented BRIEF features.
- Analysis of variance and correlation of oriented BRIEF features.
- A learning method for de-correlating BRIEF features under rotational invariance, leading to better performance in nearest-neighbor applications.

First it use FAST to find keypoints, then apply Harris corner measure to find top N points among them. It also use pyramid to produce multiscale-features.

D.oFAST: FAST Keypoint Orientation

FAST features are widely used because of their computational properties. However, FAST features do not have an orientation component. oFAST is the efficiently computed orientation added to the FAST.

FAST Detector

FAST takes one parameter, the intensity threshold between the center pixel and those in a circular ring about the center. FAST does not produce a measure of cornerness, and it has large responses along edges. Harris corner measure to order the FAST keypoints is employed. For a target number N of keypoints, it sets the threshold low enough to get more than N keypoints, then orders them according to the Harris measure, and picks the top N points. FAST does not produce multi-scale features. A scale pyramid is employed of the image, and produces FAST features (filtered by Harris) at each level in the pyramid. ORB uses a simple but effective measure of corner orientation, the *intensity centroid*. The intensity centroid assumes that a corner's intensity is offset from its center, and this vector may be used to impute an orientation. The moments of a patch can be defined as:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y)$$

and with these moments we may find the centroid:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right)$$

We can construct a vector from the corner's center, O, to the centroid, OC. The orientation of the patch then simply is:

$$\theta = \mathrm{atan2}(m_{01}, m_{10})$$

Where atan2 is the quadrant-aware version of arctan The centroid method is compared with two gradient based measures, BIN and MAX. In both cases, X and Y gradients are calculated on a smoothed image. MAX chooses the largest gradient in the key point patch; BIN forms a histogram of gradient directions at 10 degree intervals, and picks the maximum bin.

E. rBRIEF: Rotation-Aware Brief

The BRIEF descriptor is a bit string description of an image patch constructed from a set of binary intensity tests. Consider a smoothed image patch, p. A binary test τ is defined by:

$$\tau(p; x, y) = \begin{cases} 1 : p(x) < p(y) \\ 0 : p(x) < p(y) \end{cases}$$

where p(x) is the intensity of p at a point x. The feature is defined as a vector of n binary tests:

$$f_n(p) = \sum_{1 \le i \le n} 2^{i-1} \tau(p; x_i, y_i)$$

6. RESPONSES OBSERVED IN ORB

Evaluate the combination of oFAST and rBRIEF, which is called ORB, using two datasets: images with synthetic in-plane rotation and added Gaussian noise, and a real-world dataset of textured planar images captured from different viewpoints. For each reference image, the oFAST keypoints and rBRIEF features are computed, targeting 500 keypoints per image. The synthetic test set with added Gaussian noise of 10.

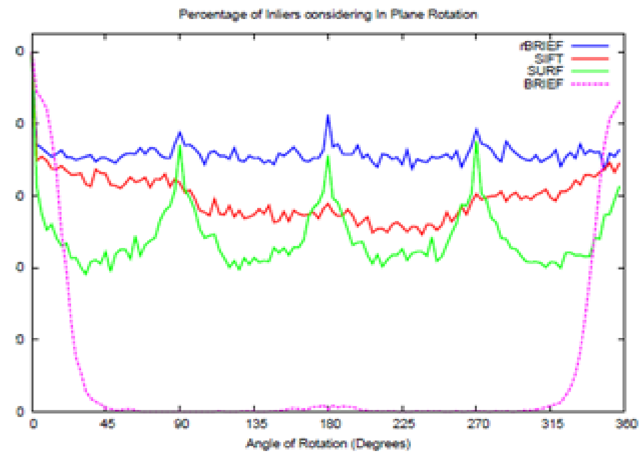

Figure 6.1: The synthetic test set with added Gaussian noise of 10.[1]

- The inlier performance vs. noise.

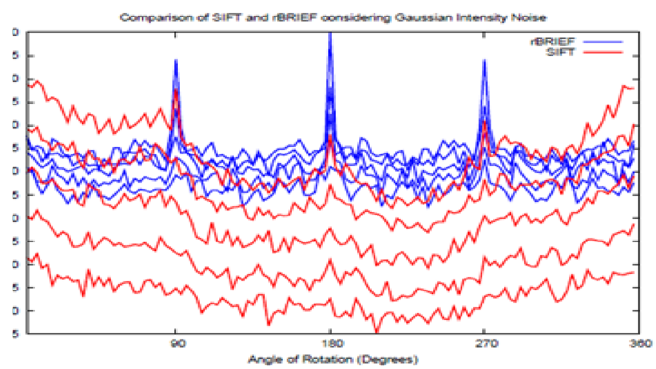

Figure 6.2: The inlier performance vs. noise.noise,at different noise levels[1]

The noise levels are 0, 5, 10, 15, 20, and 25. SIFT performance degrades rapidly, while rBRIEF is relatively unaffected. This proves the efficiency of ORB over SIFT and SURF.

7. FEATURE EXTRACTION COMPARISON

Following table gives the comparison between SIFT, SURF and ORB.

Algorithms	**Time per frame**	**Noise immunity**	**Rotation**
SIFT	Low	Moderate	Moderate
SURF	Moderate	Low	Low

ORB	High	High	high

8. CONCLUSION

As few previous studies review both image feature extraction and image feature representation, which play a crucial role in multimedia processing community. So in this paper, we provide a comparative study of the latest development in image feature extraction and image feature representation.

REFERENCES

1. Ethan Rublee, Vincent Rabaud, Kurt Konolige Gary Bradski, **“**ORB: an efficient alternativeto SIFT or SURF”**.**
2. OndrejMiksik, “Evaluation of Local Detectors and Descriptors for Fast Feature Matching".
3. Michael B. Holte, Cuong Tran, Mohan M. Trivedi, Thomas B. Moeslund,”Human Pose Estimation and Activity Recognition From Multi-View Videos: Comparative Explorations of Recent Developments".
4. Herbert Bay, TinneTuytelaars, and Luc Van Gool,”SURF: Speeded Up Robust Features".
5. Yan Ke, Rahul Sukthankar, “PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", International Journal of Conceptions on Electronics and Communication Engineering,Vol. 1, Issue. 1, Dec 2013; ISSN: 2357 2809.
6. ViniVidyadharan, and SubuSurendran, “Automatic Image Registration using SIFT-NCC”, Special Issue of International Journal of Computer Applications (0975 – 8887) , pp.29-32, June 2012.
7. D. Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”, Accepted for publication in the International Journal of Computer Vision, pp. 1-28, 2004.
8. Edward Rosten and Tom Drummond, “Machine learning for high-speed corner detection", International Conference on Computer Vision,2008.
9. Edward Rosten and Tom Drummond, “Fusing points and lines for high performance tracking",European Conference on Computer Vision,2006.
10. Dilip K. Prasad, “Survey of The Problem of Object detection In Real Images",International Journal of Image Processing(IJIP),Volume(6):Issue(6):2012.
11. Mikolajczyk, K., Schmid, “ An affine invariant interest point detector.ECCV.(2002) 128.
12. S.Winder and M. Brown, “Learning Local Image Descriptors,” *in Proc.CVPR’07*, 2007