IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

A Brief Study of Different Feature Detector and Descriptor

Nikita Kaushik[1], Ritu Rawat[2], Anshika Bhalla[3]

Student, CSE, Graphic Era University, Dehradun, India[1,2,3]

Abstract: Robust, accurate and automatic registration of image is complex task in various applications. To achieve image alignment/registration, required steps are: detection of features, Feature matching, source of transformation function. Extraction of different consistent features from image is very useful and that may be used to gain reliable matching between various views of a scene or object. In addition of deciding where and to what range a feature occur in an image, there is a distinctive body of research to decide how to represent the neighbourhood of pixels near a localized region, called feature descriptor. The easiest technique is scaled for the size of the region, use of pixel intensity values, or representation of eigen space. This survey contains various descriptors like SIFT(Scale-invariant feature transformation), SURF(Speeded up robust features), HOG(histogram of gradient), MSER.

Keywords: Key point, SIFT, Feature description, Feature extraction.

I. INTRODUCTION

A fundamental phase of computer vision is image matching, including scene or object recognition, stereo correspondence, solving of 3D structure from various images, and motion tracking. Matching of key point descriptor and recognition are two main approaches in computer vision. This survey include features of image that have various characteristics that make them appropriate for differing images matching of a scene or object. Features are invariant to scaling and image rotation, and partially consistent to viewpoint of 3D camera and change in illumination. From example images feature descriptors are draw out. These descriptors have to be different and, at the same time they are robust to noise and detect errors. After that feature descriptors are matched between distinct images. Matching of feature descriptors can be depending on Euclidean distance.

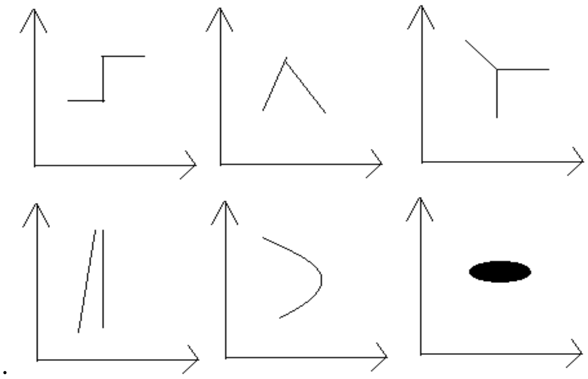

Fig 1. Types of key points, contain interest point and corners. (Left to right) Step, roof, corner, edge or line, contour or ridge, maxima region

Algorithms which are used to extract the key points may be called as detectors, and the algorithms used for describe the features may be referred s descriptors. Features may be treated a set collected of (1) corners, (2) interest points, (3) contours or edges, (4) regions or larger features like blobs.

If we select the extracted interest points carefully then it is wonted that the extracted key points will provide the appropriate information from input data. It would be useful to accomplish a wanted task by using compressed representation rather than entire size of input. Main steps of computation which are used to generate the series of image features are given below:

- Detection of scale-space extrema : The first step of computation is searching of image locations and over all scales. This is implemented appropriately with the help of DoG(difference-of-Gaussian) function to determine potential key points that are consistent to orientation and scale[1].
- Localization of key point: To determine scale and location ,a specific aspect model is adapted at every candidate location. Selected key points are depending on their measured stability.
- Orientation assignment: Every key point locations are assigned to one or more orientations which is based on gradient directions of local image. All future workings are accomplish on image input that has been converted to the assigned location, orientation, and scale for every feature, thereby supporting invariance to these transformations.
- Key point descriptor: Gradient of local images are calculated at the elected scale in the domain around every key point. These are converted into a description that acknowledge for indicative levels of change in illumination and distortion of local shape.

The features should be exclusive i.e. if same point is being explained in two or more images then that point should have same explanation and it have proper dimensions[1]. If a descriptor is large then it will create the larger computation, but a small descriptor may be reject the some useful information. Figure 2 shows the feature descriptor extracted from an image.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

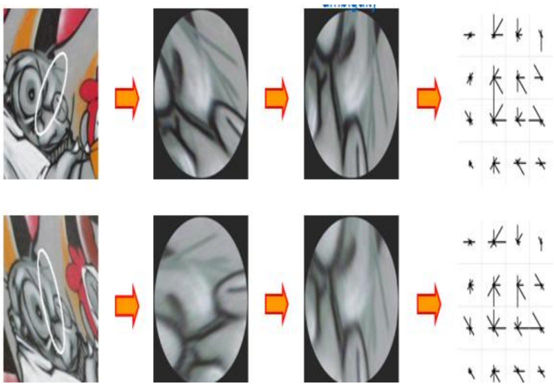

Fig 2. Extracted descriptors

II. PROPERTIES OF GOOD FEATURES

A. Repeatability
If there are 2 images of the similar scene or object, captured under distinct viewing environment, a elevated percentage of the key points extracted on the scene or object part visible in both images should be found in both images[2].

B. In formativeness/distinctiveness
For distinguish or match the features, the intensity sequences underlying the extracted key points should show a lot of changes.

C. Locality
Features should be local, so as to decrease the chances of occlusion and to allow simple model approximations of the geometric and photometric deformations between two images taken under different viewing conditions.

D. Quantity
The count of extracted key points should be adequately large, so that an average number of features are extracted even on teeny objects. However, the selected number of features relies upon the application. with the help of intuitive and simple threshold, the number of extracted features should be versatile over a large area. The denseness of features should emulate the information matter of the image to serve a compressed image representation.

E. Accuracy
The extracted features should be exactly localized, both in image location, as with respect to possibly shape and scale.

F. Efficiency
Preferably, in a new image the extraction of features should allow for time-critical applications.
Repeatability, the most valuable characteristic of all, can be attained in two distinct ways: either by robustness or invariance.

G. Invariance
When large disfigurement is to be assumed, the preferred method is to arrange these mathematically if possible, after that develop approach for feature extraction that are unaltered by these mathematical conversions[2].

H. Robustness
If there are relatively small transformations, it often sufficient to create feature extraction method less delicate to such transformations, i.e., the accurateness of the extraction may reduce, but not excessively so. Typical transformations that are attempted using robustness are discretization effects, image noise, blur, compression artefacts, etc. Also photometric and geometric aberrations from the mathematical standard used to achieve invariance are often overcome by including more robustness.

III. FEATURE DETECTOR AND DESCRIPTOR

Most of the designed features support both a detector algorithm and a descriptor-algorithm. We can combine these algorithms among each other. For a situation, the key points detected by SIFT can be described by a FREAK descriptor.. Combinations should be effective with respect to a particular application. All feature descriptor and detectors define some steps of image processing

A. Scale Invariant Feature Transformation(SIFT)
It was proposed by Lowe and gives feature that are invariant to affine distortion , scale, illumination changes, noise, rotation and 3D viewpoint changes.

SIFT algorithm has 4 main steps:
- Scale Space Extrema Detection
- Key point Localization
- Orientation Assignment
- Description Generation

The first step is to recognize the scales and location by scale space extrema in the Difference-of-Gaussian(DoG) function with distinct values of σ. Scale space separated by a constant factor M as given in the equation

$$\mathbf{D(x, y, \square\square) = (G(x, y, M\square) - G(x, y, \square\square)) \times I(x, y)} \quad \text{...(1)}$$

Where , I is image and G is the function of Gaussian.

To generate a DoG subtraction of Gaussian images are performed. Subsample of Gaussian images is done by a factor of 2 after that a DoG is generated by sampled images.

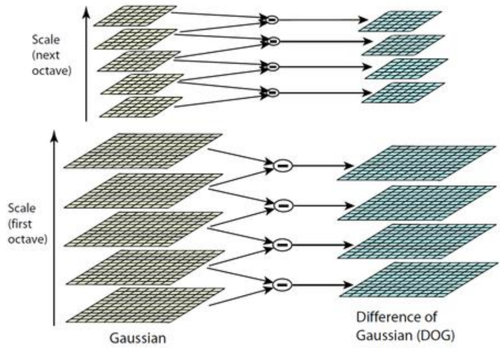

Fig 3. DoG pyramid

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

To detect the local minima and maxima of D(x, y, σ), a pixel is compared by 3*3 neighbourhood. Figure 3 shows the DoG pyramid.

In localization step, key points are localized and filtered by removing the key points where the low contrasrt points are rejected by them.

In orientation assignment stage, the obtain key point's orientation is depend on local image gradient. In description generation step local image descriptor of every key point depend on image orientation and gradient magnitude at every sample point in a domain centred at key point[3,4].

Fig 4. Original image

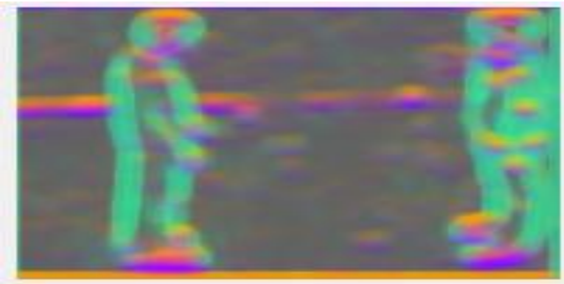

Fig 5. After apply SIFT descriptor

B) Speeded-Up Robust Feature(SURF)

It is a local featuredescriptor and detector. It is also used for object recognition,registration, 3D reconstruction, and classification. It is inspired by SIFT descriptor. it (standard version) is more faster than SIFT. It uses Haar wavelet. SURF is better in terms of speed.

- Key point localisation: detection of Blob-like structures are performed where element of the Hessian Matrix is maximum. To localize Blob-like interest point it uses a scale-space over a 3*3*3 neighbourhood. To identify feature orientation, a series of HAAR-like feature responses are calculated in local domain surrounding every interest point inside a circular radius, calculated at the matching pyramid scale for the interest point. the given point x=(x, y)in image I, the Hessian matrix H(x, σ) in x at scale σ, it can be express as

$$\mathbf{H}(\mathbf{X},\boldsymbol{\sigma}) = \begin{bmatrix} \mathbf{L}_{\mathbf{xx}}(\mathbf{x},\boldsymbol{\sigma}) & \mathbf{L}_{\mathbf{xy}}(\mathbf{x},\boldsymbol{\sigma}) \\ \mathbf{L}_{\mathbf{yx}}(\mathbf{x},\boldsymbol{\sigma}) & \mathbf{L}_{\mathbf{yy}}(\mathbf{x},\boldsymbol{\sigma}) \end{bmatrix} \qquad \ldots (2)$$

Where Lxx (**x**, σ) is the convolution result and same for Lxy (**x**, σ) and Lyy (**x**, σ).

- Key point Description: On the integral images first order Haar wavelet acknowledgements are applied in X and Y directions. After that the orientation should be regenerated depend on the information from the circular domain over the key point. The descriptor of key points are computed by a square region adjusted to the selected orientation. At the end, features are matched[5,6].

Fig 6. Original image

Fig 7. After apply SURF descriptor

C) Histogram Of Gradient(HOG)

This descriptor used in image processing and computer vision for detection of object. In loacalized ares of images this technique counts appearances of gradient orientation.it is similar to SIFT descriptor, edge orientation histogram and shape contxet , but its is differ in terms of computed dense grid of evenly spaced cells. To improve accuracy it uses overlying local contrast localization.

HOG functions on unformed data; while various methods depend on Gaussian smoothing and another filtering techniques to construct the data, HOG is created particularly to use all the unformed data without inserting filtering artefacts that delete fine details. It's a trade off: filtering artefacts like smoothing vs. image artefacts such as refined details. The HOG method represents preferential results for the raw data.

HOG taxonomy:

- Spectra: Gradient histograms of local region
- Feature shape: Circle or rectangle
- Feature pattern: Dense 64x128 typical rectangle
- Searching method: Grid over scale space
- Distance function: Euclidean
- Feature density: Dense overlapping blocks
- Robustness: 4 (noise, illumination, scale, viewpoint)

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Fig 8. Original Imge

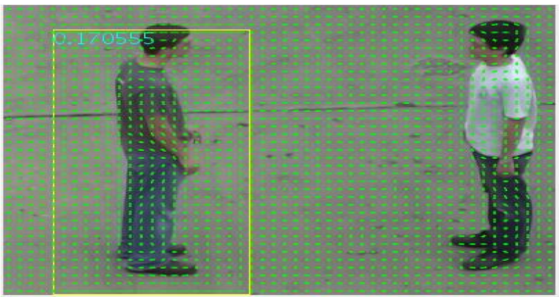

Fig 9. After apply HOG descriptor

Fig 10. Original Image


Fig 11. After apply MSER descriptor

D) Maximally stable extremal regions(MSER)

In images it is used as a blob detection method. This technique was intoduced by Matas. It is used to determine the correlation between image portions from 2 images along with various viewpoints.This technique of selecting an absolute number of correlative image elements devote to the wide-baseline matching.it also leads the better object recognition and stereo matching algorithm. Regions of MSER are related to morphological blobs and are properly robust to lighting and skewing. MSER is basically an adequate alternative of the watershed algorithm, except that the aim of MSER is to determine a range of thresholds that leave the watershed basin unchanged in size.

Smaller MSER regions are chosen for stereo applications and interconnection is take place for final correspondence. Similarity is also evaluated under a set of circular MSER regions at chosen rotation intervals. Few interesting advantages of MSER are given below:

- Multi-scale detection and multi-scale features. Since these features do not need any scale space or image smoothing, both fine-edge features and coarse features can be detected.
- Variable-size features not restricted to search window size or patch size they are calculated globally over an entire region.
- Invariance of affine transformation, which is a specific goal.
- General invariance in stability of detection, and to shape change, since the maximal regions attend to be extracted over a large range of image transformations.
- We can consider the MSER on the basis of a shape descriptor and also take as the alternative of segmentation for morphological methods. Every MSER region can be described and analyzed using shape metrics[7].

E) Binary Robust Invariant Scalable Key Points(BRISK)

It is a local binary approach . BRISK uses a circular-symmetric pattern region shape and a aggregate of 60 point-pairs as line segments arranged in 4 concentric rings[9].

BRISK takes point-pairs of both long segments and short segments, and this supports a measure of scale invariance, since coarse resolution may be mapped better by long segments and fine resolution may be mapped by short segments[10]. Algorithm of BRISK is unique .

The main computational steps of the algorithm are given below:

- In scale space key points using AGHAST or FAST based selection.
- To achieve the point value apply Gaussian smoothing at every pixel sample point.
- Makes 3 sets of pairs: Short pairs, long pairs, and unused pairs (unused pairs are not in the short pair or long pair set).
- Calculate gradient between long pairs, sums gradients to drive orientation.
- Gradient orientation are used to rotate and adjust short pairs.
- Generates binary descriptor from short pair point-wise comparisons.

F) Fast Retina key points(FREAK)

Alexandre Alahi, in their paper “FREAK”[6], discussed around novel key point descriptor influenced by human visual system and more specifically the retina, coined FREAK. A descend of binary strings is calculated by effectively analyzing image intensities around a retinal sampling pattern.

Various sampling grids are available to compare pixel Intensities of pairs. ORB and BRIEF use random pairs. A circular pattern is used by BRISK where points are evenly separated on circles concentric. For all the tested image transformation FREAK is more robust. Surprisingly, in the initial testing environment SIFT is the unfavourable descriptor, similar to what has been shown in BRISK. BRISK is slower than FREAK but BRISK is 2 orders of magnitude faster than SURF and SIFT[8].

FREAK Summary Taxonomy

- Shape of feature : Square
- Density of feature : Sparse local at AGAST key points
- Searching method: Sliding window is used over scale space
- Feature pattern: 31x31 region pixel point-pair compares
- Function of distance : Hamming distance
- Spectra: Local binary orientation vector + coarse-to-fine
- Robustness: 6 (contrast, brightness, view point, rotation, scale, blur)

G) BRIEF

This is a local binary descriptor and has gained very good performance and accuracy in robotics applications. For binary comparison BRIEF uses an arbitrary distribution pattern of 256 point-pairs in a local 31x31 region, to generate the descriptor. For comparison BRIEF select the random pairs of points in local region.

BRIEF Summary Taxonomy

- Spectra: Local binary
- Search method: Sliding window
- Feature shape: Square centred in interest point
- Feature density: Local 31x31 at interest points
- Distance function: Hamming distance
- Robustness: 2 (contrast, brightness).
- Feature pattern: Randomly comparison of local pixel point-pair

H) Oriented fast and rotated brief(ORB)

ORB is a fast descriptor. it is based on BRIEF. It is invariant to rotation and noise resistant. This feature build on the basis of recently developed and a fast key point detector BRIEF descriptor, because of this reason it is known as ORB. Both these methods are drawing attention because of their low cost and good performance[11].

Some characteristics of ORB are given below:

- The extension of an accurate and fast orientation component to FAST.
- The effective calculation of oriented BRIEF features.
- Study of correlation and variance of oriented BRIEF features.
- A learning technique for de-correlating BRIEF features inside rotational invariance, that leads better result in nearest-neighbour applications.

ORB Summary Taxonomy

- Spectra: Orientation vector + local binary
- Pattern of feature: comparison of trained local pixel point-pair
- Density of feature: Local 31x31 at interest points
- Searching method: Sliding window
- Distance function: Hamming distance
- Feature shape: Square
- Robustness: 3 (contrast, brightness, rotation, *limited scale)

IV. CONCLUSION

In this survey, feature detector and descriptor were first popularized in case of their use in application of computer vision. The survey describes various efficient techniques of key point extraction. Key points supports the capability to identify image correspondences despite of change in view conditions ,occlusion, or the presence of clutter. Here we explained various methods for effective key point detection, but every descriptor have its own demerits and merits. According to the current requirement ,we can select our descriptor. A comparative study and a brief summary of every key point descriptor method is provide here.

ACKNOWLEDGMENT

We would like to thank all our friends, guides and family members who provided support and followed up with us for such a long period of time.

REFERENCES

[1] D. Lowe, Distinctive Image Features from Scale-Invariant Keypoints , Accepted for publication in the International Journal of Computer Vision, pp. 1-28, 2004.

[2] Tinne Tuytelaars and Krystian Mikolajczyk, Local Invariant Feature Detectors: A Survey, Foundations and Trends in Computer Graphics and Vision Vol. 3, No. 3 (2007) 177–280.

[3] 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.

[4] P M Panchal, S R Panchal, S K Shah, A Comparison of SIFT and SURF, International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 2, April 2013.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, Speededup robust features (surf), Comput. Vis. Image Underst., vol. 110, no. 3, pp. 346–359, June 2008.

[6] Alexandra Derntl, Survey of Feature Detectors and Descriptors in Surgical Domain, 2014.

[7] Rekhil M Kumar, Sreekumar K, A Survey on Image Feature Descriptors , International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 7668-7673

[8] A. Alahi, R. Ortiz, and P. Vandergheynst, FREAK: Fast Retina Keypoint, in IEEE Conference on Computer Vision and Pattern Recognition, ser. IEEE Conference on Computer Vision and Pattern Recognition. New York: Ieee, 2012, cVPR 2012 Open Source Award Winner.

[9] M.Calonder, V. Lepetit, C. Strecha, and P. Fua, Brief: Binary robust independent elementary feature, . In *In European Conference on Computer Vision*, 2010.

[10] Pami Prakash, Anil A.R, Survey on Key Feature Descriptors Used In Computer Vision Applications, International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2, Issue-1, January 2016.

[11] Ethan Rublee, Vincent Rabaud, Kurt Konolige Gary ,ORB: an efficient alternative to SIFT or SURF.