INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

Local Feature based descriptors and their applications

Ridhi Jindal, Dr. Sonia Vatta
ridhijindal136@gmail.com

Abstract: *This paper presents a study on SIFT (Scale Invariant Feature transform) which is a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection and SURF (Speeded-up Robust features) which is speeded up the SIFT's detection process without scarifying the quality of the detected points. SURF approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster.*

KEYWORDS*: SIFT,SURF, Keypoints, Scale, Descriptor, DoG.*

1. Introduction

Feature detection and image matching represent many important tasks in real world applications. Their application continues to grow in a variety of fields day by day. From simple photogrammetry tasks such as feature recognition, to the development of sophisticated 3D modeling software, there are several applications where image matching algorithms play an important role. Moreover, this has been a very active area of research in the recent decades and as indicated by the tremendous amount of work and documentation published around this. As needs change and become more demanding, researches are encouraged to develop new technologies in order to fulfill these needs.

Here we are giving a brief about two well-known feature detection algorithms SIFT [1] and SURF [13] and their various applications.

2. Literature Review

Robust feature detection, image matching and 3D models are concepts that have been around for many years now in the computer vision field. But it wasn't until the end of the last decade and the beginning of this one that the problem was really approached by numerous researchers and professionals working in this field. It is well known that achieving true invariant object recognition has been one of the most important challenges in computer vision. Recently, there has been a significant progress in the use and implementation of algorithms towards the detection of invariant features in every-day more complex images (Lowe, 1999).

Few years after his first publication on feature detection for textured images, Lowe published an improved version of his work and presented his results with the publication of the Scale Invariant Feature Transform (SIFT) algorithm

(Lowe, 2004) which was followed by many other algorithms like Principal Component Analysis-SIFT (PCA-SIFT) and the Speed-Up Robust Features (SURF) in 2006.

3. Scale Invariant Feature Transform: SIFT[2]

SIFT, as mentioned before, was developed by David Lowe in 2004 as a continuation of his previous work on invariant feature detection (Lowe, 1999), and it presents a method for detecting distinctive invariant features from images that can be later used to perform reliable matching between different views of an object or scene.

Two key concepts are used in this definition: distinctive invariant features and reliable matching. The approaches which Lowe's has named SIFT, and are broken down into four major computational stages [3,4]:

a) Scale-Space extrema detection

b) Keypoint localization

c) Orientation assignment

d) Keypoint descriptor

a) Scale-space extrema detection [5]

The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of Gaussian function to identify potential interest points that are invariant to scale and orientation.

b) Key point localization [6]

At each candidate location, a detailed model is fit to determine location and scale. Key points are selected based on measures of their stability.

c) Orientation assignment

One or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

d) Key point descriptor

The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

The under given is the SIFT Feature detection algorithm.

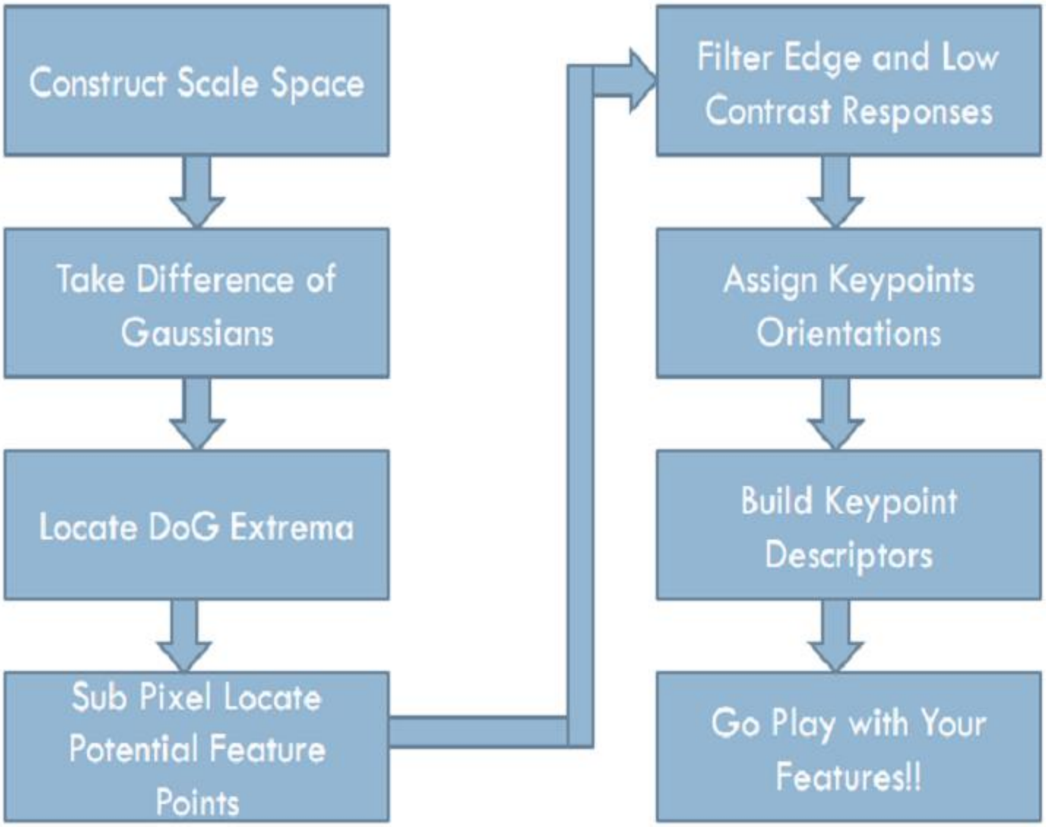

Fig: SIFT feature detection Algorithm

4. Speeded-Up Robust Feature: SURF

The Speed-Up Robust Feature detector (SURF)[14] was conceived to ensure high speed in three of the feature detection steps: detection, description and matching (Bay et al., 2006). SURF speeded up the SIFT's detection process without scarifying the quality of the detected points. SURF approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster [12]. This is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors (specially, 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. [15]

SURF detector is based on the determinant of the Hessian matrix. In order to motivate the use of the Hessian, we consider a continuous function of two variables such that the value of the function at (x; y) is given by f(x; y). The Hessian matrix, H, is the matrix of partial derivatives of the function f

$$H(f(x,y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

The determinant of this matrix, known as the discriminant, is calculated by:

$$det(H) = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

The value of the discriminant is used to classify the maxima and minima of the function y the second order derivative test. Since the determinant is the product of eigenvalues of the Hessian we can classify the points based on the sign of the result. If the determinant is negative then the eigenvalues have different signs and hence the point is not a local extremum; if it is positive then either both eigenvalues are positive or both are negative and in either case the point is classified as an extremum. [16]

5. Applications

I. Object recognition using SIFT features[5]
II. Robot localization and mapping [7]
III. Panorama stitching [8]
IV. 3D scene modeling, recognition and tracking [9]
V. 3D SIFT-like descriptors for human action recognition [10, 11]
VI. Analyzing the Human Brain in 3D Magnetic Resonance Images [16]

6. Conclusion

Both SIFT and SURF are well known Feature detection algorithms with so many applications in the existing world and has given the world of computer vision a new shape. SIFT is arguably the most popular algorithm that can match under different scales, rotations and lighting, but it was significantly slow. Many implementations can be found as open source codes in the web. SURF is a fast and performant interest point detection-description scheme which outperforms the current state-of-the art, both in speed and accuracy. The descriptor is easily extendable for the description of affine invariant regions.

7. REFRENCES

[1]. Lowe, David G. (1999). "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision 2. pp. 1150–1157.

[2]. "Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image", David Lowe's patent for the SIFT algorithm, March 23, 2004.

[3]. Lowe, D. G., “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.

[4]. Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G., Poggio, T., “A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex”, Computer Science and Artificial Intelligence Laboratory Technical Report, December 19, 2005 MIT-CSAIL-TR-2005-082.

[5]. Beis, J., and Lowe, D.G “Shape indexing using approximate nearest-neighbour search in high-dimensional spaces”, Conference on Computer Vision and Pattern Recognition, Puerto Rico, 1997, pp. 1000–1006.

[6]. Lowe, D.G., Local feature view clustering for 3D object recognition. IEEE Conference on Computer Vision and Pattern Recognition,Kauai, Hawaii, 2001, pp. 682-688.

[7]. Lowe, David G.; Little, J. (2001). "Vision-based mobile robot localization and mapping using scale-invariant features". Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) 2. p. 2051

[8]. Lowe, David G. (2003). "Recognising Panoramas". Proceedings of the ninth IEEE International Conference on Computer Vision 2. pp. 1218–1225.

[9]. Flitton, G.T., Breckon, T.P., Megherbi, N. (2013). "A Comparison of 3D Interest Point Descriptors with Application to Airport Baggage Object Detection in Complex CT Imagery". Pattern Recognition (Elsevier).

[10]. Niebles, J. C. Wang, H. and Li, Fei-Fei (2006). "Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words". Proceedings of the British Machine Vision Conference (BMVC). Edinburgh. Retrieved 2008-08-20.

[11] H. Bay. From Wide-baseline Point and Line Correspondencesto 3D. PhD thesis, ETH Zurich, 2006.

[12] H. Bay, B. Fasel, and L. van Gool. Interactive museum guide:Fast and robust recognition of museum objects. In Proceedings of the rst international workshop on mobile vision, May 2006.

[13] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded up robust features. In ECCV, 2006.

[14] M. Brown and D. Lowe. Invariant features from interest point groups. In BMVC, 2002.

[15] G. Carneiro and A.D. Jepson. Multi-scale phase-based local features. In CVPR (1), pages 736 { 743, 2003.

[16] P. C. Cattin, H. Bay, L. Van Gool, and G. Szekely. Retina mosaicing using local features. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), October 2006. in press.
.