ISSN: 0067-2904

Survey of Scale-invariant Feature Transform Algorithm

Eklas khlaf Egbash, Suha Mohammad*
Department of Computer Science , University of Technology , Baghdad , Iraq.

Abstract
The effectiveness of detecting and matching of image features using multiple views of a specified scene using dynamic scene analysis is considered to be a critical first step for many applications in computer vision image processing. The Scale invariant feature transform (SIFT) can be applied very successfully of typical images captured by a digital camera.

In this paper, firstly the SIFT and its variants are systematically analyzed. Then, the performances are evaluated in many situations: change in rotation, change in blurs, change in scale and change in illumination. The outcome results show that each algorithm has its advantages when compared with other algorithms. .

Keywords: Feature Detection, Feature Descriptor, Matching.

تحويل صفة صورة غير مرتبط بمقياس

اخلاص خلف اكباش، سها محمد صالح*
قسم علوم الحاسبات، الجامعة التكنلوجيه، بغداد ،العراق.

الخلاصة
كفاءة و فعالية استكشاف و مطابقة خواص الصورة عن طريق استخدام اكثر من منظار لمشهد معين في التحليل الدايناميكي للمناظر يعتبر خطوة اولية و اساسية لعدد واسع جدا من تطبيقات الرؤيا بالحاسوب في مجال المعالجة الصورية. تقنية تحويل الخصائص غير المعتمد على التوجيه قد تم تطبيقها على صور متعددة ماخوذة بالكاميرات الرقمية.

في هذا البحث، اول خطوة تتمثل في استخدام متغيرات هذه الخوارزمية و تحليلها. و من ثم يتم تقييم الاداء لهذه التقنية في عدد كبير من الحالات (مثل تغيير الدوران، تغيير السطوع، تغيير الموقع، تغيير الاضاءة، الخ) المعطيات النهائية بينت ان كل خوارزمية لها افضليات معينة مقارنة بالخوارزميات الباقية.

Introduction

The Scale invariant feature transform (SIFT) is normally used for extracting features from images which have stable behavior in image rotation, image translation, image illumination, image scaling and different camera viewpoint. SIFT algorithm is one of the most widely used algorithms that is mainly used for object recognition, and it is applied in many real world applications (like objects recognition and face recognition [1], SIFT input is N × N image and will produce a set of features, SIFT is invariant for changes in the camera viewpoint and illumination, its algorithm has four major steps :

1. **Extrema Detection**: is used to examine the image, the examination will be in various octaves and scales which will isolate specific points in the picture which usually differ from surroundings points. These points, will call (**extrema**), and it is the potential candidates for image features.

*Email: suhaa1974_2016@yahoo.com

For extrema detection the function of Gaussian,then considered (Figure-1) the possible function of scale space.

The Scale space of specific image can be calculated as follows,

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

The following data identify the equation identifiers:

- The input image is represented by I (x, y).
- G(x, y, σ) is the variable of scale Gaussian is which can be founded as:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{(-x^2+y^2)/2\sigma^2}$$

DOG (Difference of Gaussian) function will be convolved via convolution operation to the input image that is used for detecting the locations of the stable keypoint.

2. **Keypoint Localization:** this step is considered the Keypoint Detection, that starts with the extrema, and selecting extrema points (some of these points) to be keypoints, which is chosen from a set of candidates, by rejecting all extrema points caused by low-contrast points and by edges of the picture, when keypoints are detected, the steps to performing the fitted details to the adjoining image data for scaling, locating and finding ratio of the principal curvatures, by performing this step the low contrast points will be rejected since they will be very sensitive to noise and hard to be localized.

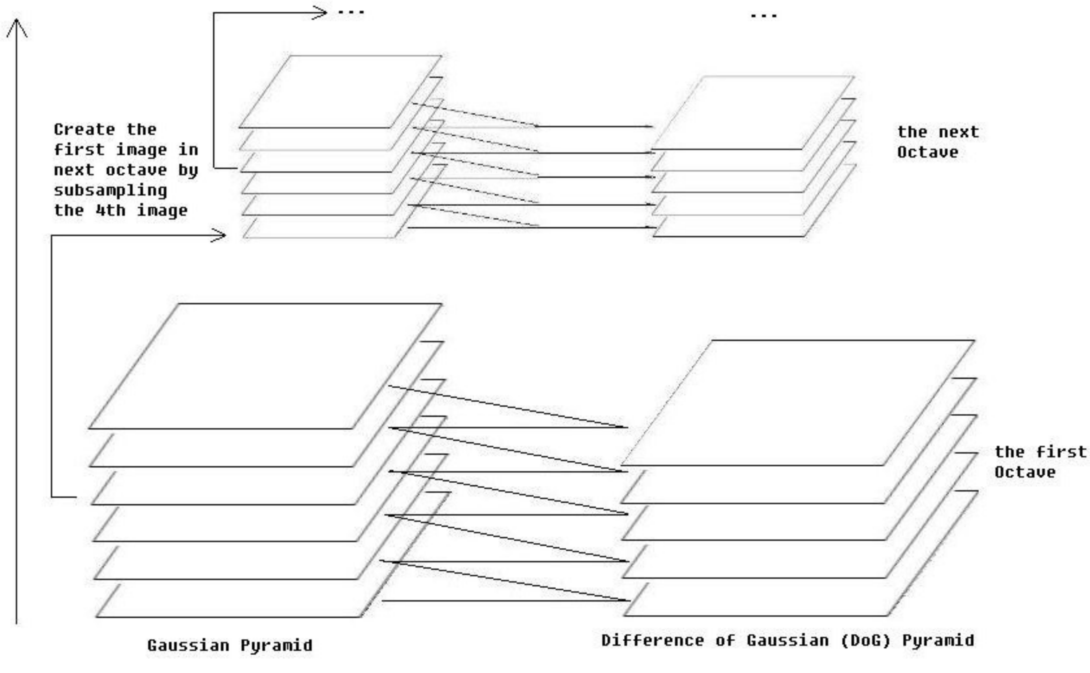

Figure 1- Gaussian function

3. **Orientation Assignment:** this step convert, the keypoint and neighborhood of these keypoints into a set of vectors, this transformation is done by computing a direction and a magnitude for each keypoint. And identify any possible keypoints that may have been missed in the first two steps; usually the points have significant magnitudes and it is not an extrema. After this step the final set of keypoints is ready.

All descriptors are represented relatively to the orientation, the SIFT will use gradient orientation histogram to determine the orientation of keypoint; this histogram will work on neighborhood of the selected keypoint. The Gaussian function will smooth the image L ,selection is made to the scale of keypoint.

m(x,y)=$\sqrt{L(x+1,y)-L(x_1,y))2+(L(x,y+1)-L(x,y_1))2}$

θ(x,y)= $\tan^{-1}(\frac{(L(x,y+1)-L(x,y_1))}{(L(x+1,y)-L(x-1,y))})$

Keypoint Descriptor Generation: The final step, the generation of Keypoint Descriptor in this algorithm, will take the predefined vector (or vectors) that are neighbors of every keypoint ,these information will consolidate to a group of 8 vectors that are called the (descriptor). Every descriptor will be convert and saved as a feature and this is done by the normalization sum of these vectors, Figure-1 shows Major phases of the SIFT_algorithm, the describer will contain 128 elements founded by (4 × 4 × 8) and the final descriptor vector is calculated by the inverting to illumination, scaling and rotation[2].

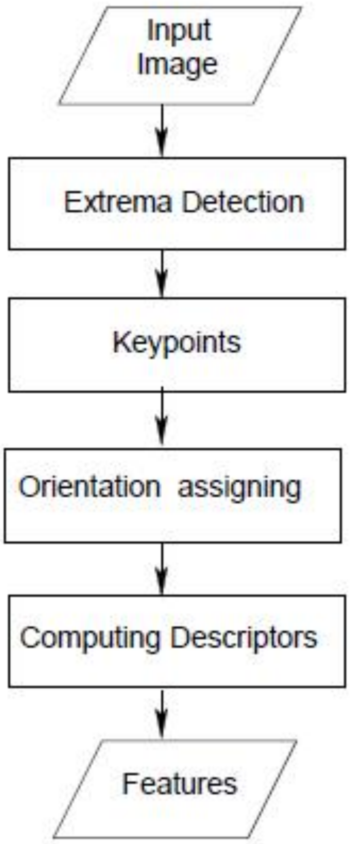

Figure 2- Major phases of the SIFT_algorithm

Scale-invariant Feature Transform Variants

In computer vision and image processing image matching is an important research field. Many problems consider it as a necessary precondition for solving many known problems, many researches improving the SIFT are dedicated to the performance in techniques of image matching, a wide variety of algorithms has been suggested [3],mainly the methods of image matching are divided into two types: matching algorithms based on local features and matching algorithms based on global. As compared with matching algorithms based on global, matching algorithms based on local feature are considered more stable. Since it’s applied very wide and in a successful manner in many real-world programs applications, as (robot localization, video data mining ,texture recognition, , retrieval of image, building panoramas, object recognition, and object category recognition)[4].

Matching algorithms based on local feature include two stages:

1. Detection of interest point.

2. Description of interest point.

Special characteristics should be taken to obtain good local features.

- Feature detection should have a high speed and high repeatability rate.

- Feature description should have a low feature dimension, which can be met easily the rotation Robustness to illumination, fast matching and the changes in view point.

David G. Lowe proposed a (Scale-invariant feature Transform) SIFT_algorithm that works with local feature description [5], [6] by using the analytical of existing invariance based feature detection methods and based on this analysis. The SIFT has accepted invariance and stability. The result detects local keypoints, and these keypoints contain a big amount of informati [6].

Figure 3- set of Leuven images with invariance in illumination.

Figures- 4, 5, 6, 7 and 8. Will take the first ten matching Keypoints and show the details between the dark image and the bright image between these images that are shown in Figure-3 to SIFT_algorithm and other variant types of SIFT, respectively).

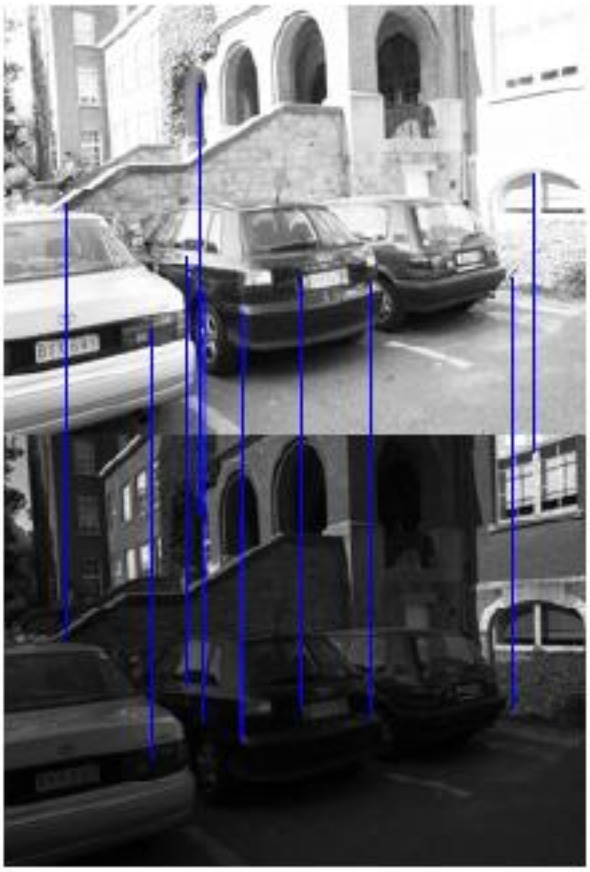

Figure 4- SIFT (shows the details of the first 10 matching

keypoints between the darkest image and the brightest image using SIFT

Because the Scale-invariant Feature Transform algorithm was proposed formally, researchers never stop improving it. Between SIFT_algorithm and its different variants, many numbers of algorithms proposed of PCA-SIFT_algorithm [7], CSIFT_algorithm [8], SURF_algorithm [9] and ASIFT_algorithm [10].

PCA-SIFT (Principle Component Analysis) [7]

This type of SIFT improvement is in the phase of descriptor establishing, in SIFT (Scale-invariant Feature Transform) the algorithm only describes the information that is considered local and doesn't use the global information, PCA-SIFT uses all the global information that is available.

PCA-SIFT considered to be strong and used widely for image dimensionality reduction data technique. Its main work is to convert a random vector (original vector), which has components that have strong correlations to a vector (new random vector), with other components with no correlations, by a perpendicular transformation. Principle Component Analysis uses the PCA for replacing the gradient histogram method in original SIFT_algorithm. PCA-SIFT description step is divided into two sub stages: descriptor establishing and generating the projection matrix. This will make the new vector, much smaller in size than a vector of standard algorithm (SIFT_algorithm).

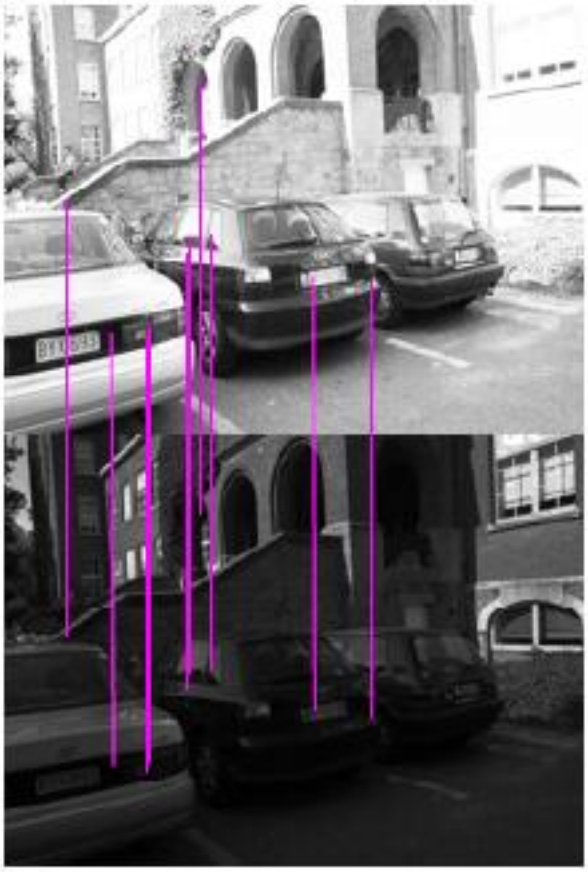

Figure 5- PCA-SIFT (shows the details of the first 10 matching keypoints between the darkest image and the brightest image using PCA-SIFT)

CSIFT (A SIFT Descriptor with Color Invariant Characteristics) [8]

This improvement will add color invariance to the basis of SIFT-algorithm (Scale-invariant Feature Transform) and plan to overcome the shortcomings of SIFT algorithm in color images.

CSIFT tries to integrate the invariance in color into the basics of original algorithm (SIFT_algorithm). The invariance in color will describe potential characteristics of optical radiation of the object in image based on a (Kubelka-Munk) theory.

Figure 6-CSIFT (shows the details of the first 10 matching keypoints between the darkest image and the brightest image using CSIFT)

SURF (Speeded Up Robust Features) [9]

This improvement is very similar to SIFT (Scale-invariant Feature Transform) but it performs different processing methods in each step, it is considered that SURF is an enhanced version of Scale-invariant Feature Transform.

The main idea of SURF algorithm is considered very similar to the original algorithm (SIFT_algorithm), different methods are use for descriptor generation and location detection. Every image in image database has a large amount of data, SURF tried to improve it by improving the description and detection effectively of extrema points. In this algorithm, for detection a quick matrix is adopted (Hessian), which has good results accuracy and speed.

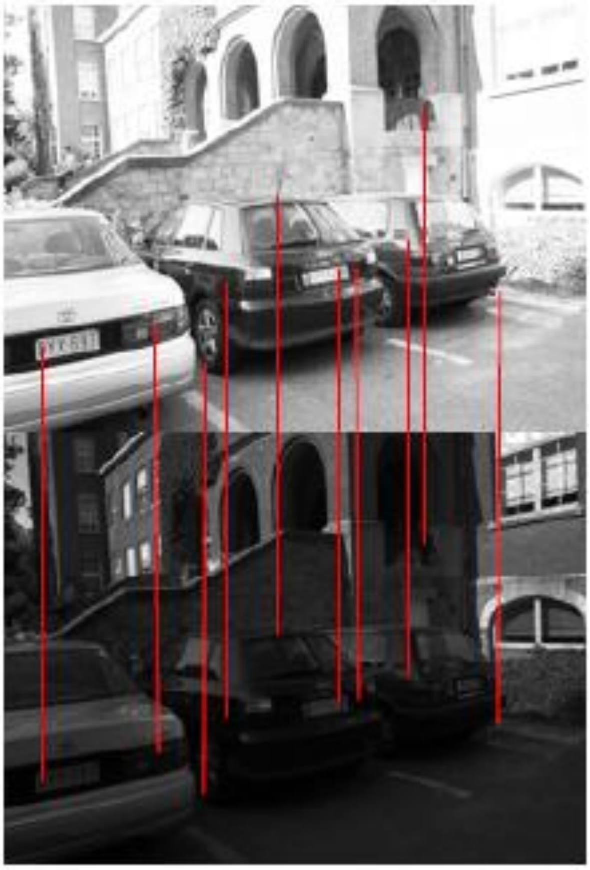

Figure 7- SURF (shows the details of the first 10 matching keypoints between the darkest image and the brightest image using SURF)

ASIFT (Affine-SIFT) [10]

This improvement is used to follow the affine transformations parameters for correcting the images and intending to resist strong affine issues. SIFT doesn’t work very well with images that have affine changes. SIFT will simulate rotation the optical axis of camera. Original algorithm will adopt a model for the transformation of image affine which will result from changing the viewpoint, the fast movement of the camera which causes deformation to the image of the measured calculated object. The angle, which can be produced by mapping camera plane optical axis and object normal plane that's measured is defined as a longitude angle. ASIFT first rotation transformation will be added to images, and secondly will detect the keypoints; lastly it will establish the affine image description.

Figure 8- ASIFT (shows the details of the first 10 matching keypoints between the darkest image and the brightest image using ASIFT)

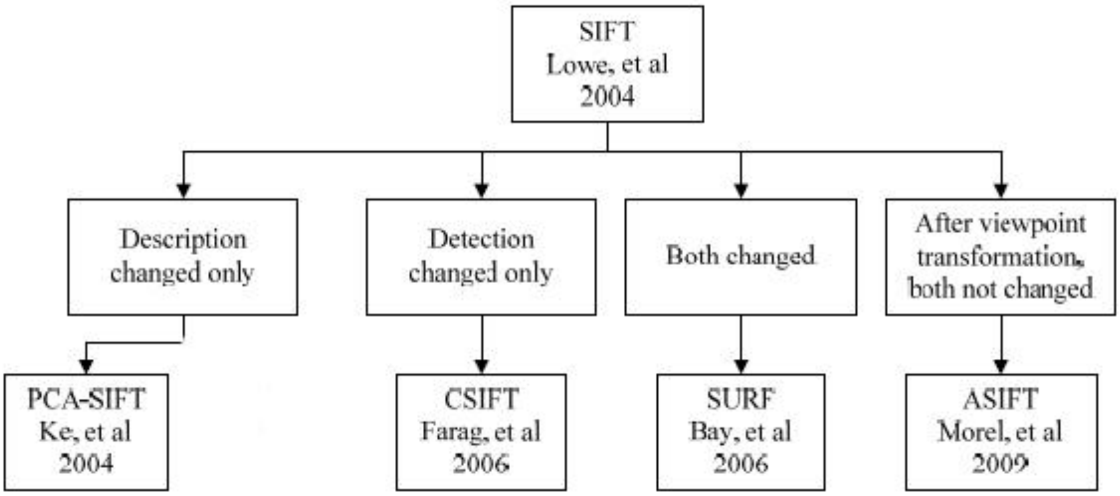

Figure 9- Relationship between SIFT and its variants.

Another Modified SIFT Algorithms

Many algorithms are published as a modification of the original SIFT_algorithm, by replacing and adding special transformation to it , mainly these algorithm are used for keypoint description .

F-SIFT (Fourier Transform SIFT)[11]

SIFT_algorithm uses the gradient data for extracting, which may not be the excellent methods in particular medical images, SIFT_algorithm mainly suffered from the false pairing problem, the solution to this problem is to extract features that are suitable from medical purposes. Using the Fourier transform is one of the solutions to this problem since its ability for describing function that considered as periodic (e.g. repetitive patterns inside texture). Further, SIFT_algorithm is strong to different illumination, and transformations. F-SIFT will use the mix properties of SIFT and Fourier. Keypoints are obtained by the SIFT_algorithm and these keypoints will enter the Fourier transform for descriptor these keypoints.

S-SIFT (S Transform SIFT)[11]

SIFT_algorithm uses the gradient data for extracting which may not be the excellent methods in particular medical images, SIFT_algorithm mainly suffered from the false pairing problem, the solution to this problem is to extract features that are suitable from medical purposes. Using the S-transform is one of the solutions to this problem due to its ability for describing function that considered as periodic (e.g. repetitive patterns inside texture). Further, SIFT_algorithm is strong to different illuminations, and transformations. S-SIFT will use the mix properties of SIFT and Fourier. Keypoints are obtained by the SIFT_algorithm and these keypoints will enter the Fourier transform for descriptor these keypoints.

Conclusions

This paper showed the SIFT family major members, which are (SIFT, CSIFT , PCA-SIFT, ASIFT and SURF), these algorithms are used for description of local feature based on scale space, evaluated the performance of these algorithm according to (blur change, scale change , rotation change, affine change, and illumination change) and time computing.

After studying each algorithm, we concluded the following results:

- In terms of rotation and scale variations, SIFT and CSIFT perform better than other algorithms, CSIFT improves SIFT in the case of changes in affine and blur, but not changes in illumination.
- ASIFT performs the best among other algorithms in the cases of affine changing, dynamic scenes and illumination of the image.
- PCA-SIFT are in the second place in terms of performance when compared to the others.
- SURF is the fastest algorithm among other SIFT variants, but its results are the worst compared to the others in terms of detection.
- Many other types of the SIFT are developed that are same as the original algorithm and only replace the of keypoint description.

References

1. Lowe, D. G. **2004**. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, **60**(20): 91-110.
2. Vinukonda, P. **2011**. A Study of the Scale-Invariant Feature Transform on a Parallel Pipeline. *Diss. Louisiana State University*, pp: 5-14.
3. Ouyang, W. **2012**. Performance evaluation of full search equivalent pattern matching algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, **34**(1): 127-143.
4. Kim, D ., Rho, S.and Hwang, E. **2012**. Local feature base multi-object recognition scheme for surveillance. *Engineering Applications of Artificial Intelligence*, **25** (7): 1373–1380.
5. Lowe, D. G. **1999**. Object recognition from local scale-invariant features. Computer vision, The proceedings of the seventh IEEE international conference on Computer Vision, Corfu, Greece, pp: 1150-1157.
6. Lowe, David G. **2004.** Distinctive image features from scale-invariant keypoints. *International journal of computer* vision, **60** (2): 91-110 .
7. Ke, Y. and Rahul, S. **2004**. PCA-SIFT: A more distinctive representation for local image descriptors. Computer Vision and Pattern Recognition,. Proceedings of the IEEE Computer Society Conference on, **2**,pp: 506-513.
8. Abdel-Hakim, A. E. and Farag, A. A.. **2006.** CSIFT: A SIFT descriptor with color invariant characteristics. Computer Vision and Pattern Recognition, *IEEE* Computer Society Conference on, **2**, pp: 17-22.
9. Bay, H. , Tuytelaars, T. and Gool, L.V. **2006**. *Surf: Speeded up robust features*. Computer vision. ECCV. Springer Berlin Heidelberg, pp: 7-13.
10. Morel, J.M. and Yu, G. **2009**. ASIFT: A new framework for fully affine invariant image comparison. SIAM *Journal on Imaging Sciences*, **2** (2).
11. Mehrotra, H., Majhi, B. and Sa, P.K. **2011**. Unconstrained iris recognition using f-sift. In Information. Communications and Signal Processing (ICICS) 2011 8th International Conference on, pp**: 1-5.**