Study of Image Feature Matching Algorithms: SIFT and SURF

R.M.Mulajkar, V.V.Gohokar,D S Galhe

R.M.Mulajkar is is working as Assisatnt Professor in Department of E & TC, JCOE, Kuran, Pune M.S., India (e-mail: rahul.mulajkar@gmail.com).
V.V.Gohokar, is working as Associate Professor in Department of E & TC, M.I.T., Pune.M.S.India (e-mail: vvgohokar@rediffmail.com).
D S Galhe is is working as Assisatnt Professor at JCOE, Kuran, Pune M.S., India

***Abstract*— There have been advances in object recognitition and image matching through matching invariant image features in the recent past. The SIFT and SURF algorithms have become more popular. Scale Invariant feature transform (SIFT) is an algorithm in image processing to detect scale-invariant feature. The sale-invariant features are the features that are detectable from images which are invariant to image scaling, rotation, affine changes and partially invariant to illumination changes. The number of improvement has been made on SIFT. The new algorithms have been developed are PCA-SIFT, GSIFT, CSIFT, SURF and ASIFT.The Speeded up robust features (SURF) algorithm uses various methods for location detection and description generation. In this paper we study the SIFT and SURF algorithm in details and evaluate performance of SIFT in different condition: Scale change, rotation change, intensity change.**

Index Terms— Interest points, SIFT, SURF, Image Matching.

I. Introduction

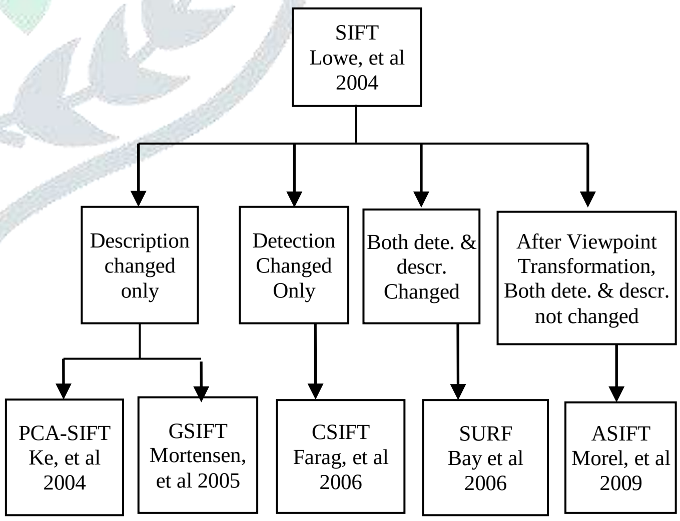
Image matching is an important step in the application of computer vision and image processing. The numbers of research are performed for image matching. Matching of image to establish their similarity is a major problem in computer vision and image processing applications. In general image matching algorithms can be classified into two groups: global and local feature based algorithm. Global feature based algorithm aims at recognizing an object as whole. To achieve matching after acquisition of object, the test object is sequentially pre-processed and segmented and then global features are extracted. Global feature based algorithm is suitable for recognition of texture less object which can be easily segmented from image background. The global feature based algorithm is simple and fast but there is drawback of changing the features with respect to illumination and scale. Thus, local feature based algorithm is mostly used and it is invariant to illumination and scale [7].The local feature based algorithm involves following steps: The first step is the extraction of salient feature points from both test and train images. The second stage is construction of region around the salient features. The final stage is matching between train and test images depending on extracted invariant features. David Lowe proposed local feature description algorithm called SIFT in 1999.SIFT has good stability and transformation. It detects local key point which contains large number of information.

The Paper is organized as follows: In section II, we review the stage in the development of matching image algorithm. In section III, we describe SIFT algorithm and SURF algorithm for image matching in detail. we conclude the paper in section IV.

II. Related Work

To extract distinct invariant features from image, SIFT algorithm is developed. From SIFT algorithm the number of researchers have worked on image matching such as PCA-SIFT, GSIFT, CSIFT, SURF and ASIFT [03]-[06].Each algorithm has its own advantages. SIFT and CSIFT performs best under scale & rotation change. CSIFT improves SIFT under noise change but not illumination change. GSIFT perform the best under the blur and illumination change. ASIFT performs best under affine change. SURF performs badly in different situations but runs fastest.

The classification of various SIFTs can be summarized as follows [16]:



The characteristics of various SIFT can be summarized as in Table 1.

Table 1.Characterstics of various SIFTs

Algorithm	Abbreviation	Keypoint Detection		Key point Description		
		Scale-space	Selection	Main direction	Feature Extraction	#Dimensions
SIFT	Scale invariant feature transform	Different-scale images convoluted with a Gaussian function	Detect extrema in DoG space; do non-maxima suppression	Calculate a gradient amplitude of a square area; regard the direction with the maximum gradient	Divide a 16×16 region into 4×4 sub-regions; create a gradient histogram for each sub-region	128
PCA-SIFT	Principal Component Analysis-SIFT	Same as SIFT	Same as SIFT	Same as SIFT	Extract a 41×41 patch; form a 3042-dimension vector; use a project matrix	20 or less
GSIFT	Global Information SIFT	Same as SIFT	Same as SIFT	Same as SIFT	For each keypoint, create a vector consisting of SIFT description and a global texture vector	188
CSIFT	Color invariance-SIFT	Replace grayscale with color invariant; convolute with	Same as SIFT	Same as SIFT	Same as SIFT	384
SURF	Speeded-up Robust Feature	Different-scale box filter convoluting with an original image	Use a Hessian matrix to determine candidate keypoints; do non-maxima suppression	Calculate a Haar wavelet output in both x and y directions of each sector in a circular area	Divide a 20×20s region into 4×4s sub-regions; calculate a Haar wavelet response	64
ASIFT	Affine-SIFT	After a preprocessing - viewpoint transformation, follow SIFT's steps (i.e., the same as SIFT)				

III. Analysis of Image Matching Algorithms

A. Scale-Invariant Feature Transform:

We conducted experiment to study the performance of SIFT and its variation effect under different situation: scale, rotation, blur, illumination and affine change. SIFT algorithm is implemented in Matlab 2015a and executed on Dell Laptop. It has processor Intel (R) Core i3 CPU,M380@2.53GHZ,RAM 4GB, 64 bit windows operating system.

Geometric transformation modifies the spatial relationship between pixels in an image. These transformations often are called rubber-sheet transformation because they may be viewed as analogous to 'printing' image on sheet of rubber and then stretching sheet according to predefined set of rules. In terms of digital image processing, geometric transformation consists of two basic operations: spatial transformation of coordinates and intensity interpolation that assigns intensity values to spatially transformed pixels.

The transformation coordinated may be expressed as

$$(x', y') = T\{s_x, s_y\}$$

Where (s_x, s_y)-pixel coordinate in the original image and (x', y')-corresponding pixel in transformed image.

This transformation can scale, rotate, transfer or sheer set of coordinates

$$\begin{bmatrix}x'\\y'\\1\end{bmatrix}=\begin{bmatrix}1&0&tx\\0&1&ty\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

Translate

$$\begin{bmatrix}x'\\y'\\1\end{bmatrix}=\begin{bmatrix}Sx&0&0\\0&Sy&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

Scale

$$\begin{bmatrix}x'\\y'\\1\end{bmatrix}=\begin{bmatrix}cos\theta&-sin\theta&0\\sin\theta&cos\theta&0\\0&0&1\end{bmatrix}$$

Rotate

$$\begin{bmatrix}x\\y\\1\end{bmatrix}\begin{bmatrix}x'\\y'\\1\end{bmatrix}=\begin{bmatrix}a&b&c\\d&e&f\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

Affine

To analyze algorithm SIFT two images are taken as experimental data as shown in figure

Fig.1. a) Image1 b) Image2

In scale-Invariant Feature Transform [SIFT] following are the major steps performed

1. Scale-space extrema detection 2.Keypoint localization

3. Orientation assignment 4.Keypoint description

1. Scale-space extrema detection:

In this step detecting location that are invariant to scale change of image can be achieved by searching for stable features across all scale scale-space. The scale-space function is defined as function L(x, y, σ) where σ is variance, produced from convolution of Gaussian G(x, y, σ) with input I(x, y).

Mathematically it can be given as-

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

Where *-convolution operator in x & y

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

This method is used to detect stable keypoint i.e.it uses scale-space extrema in the difference of Gaussian (DoG) function convolved with image D ((x, y, σ) which can be determined by differences of two adjacent scales separated by constant k can be given as

$$D((x,y,\sigma) = \{ G(x,y,k\sigma)- G(x,y,\sigma)\} * I(x,y)$$
$$= L(x, y, k\sigma) - L(x, y, \sigma) \quad (2)$$

The relation between D and Laplacian of Gaussian σ^2G is given as-

$$\frac{\partial G}{\partial \sigma} = \sigma\ .G$$

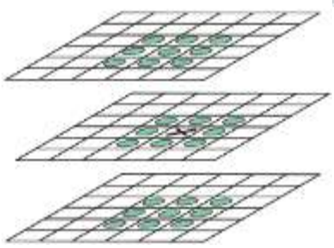
$$\sigma\ .G = \frac{\partial G}{\partial \sigma} = \frac{G(x,y,k\sigma) - G(x,y,\sigma)}{(K\sigma - \sigma)}$$

$$G(x, y, \sigma) - G(x, y, \sigma) = \sigma\ .G(K\sigma - \sigma)$$
$$= \sigma\ .G\ (k-1) \quad (3)$$

In our experiment we take $k=\sqrt{2} = 1.414$ as there is no impact on stability of extrema detection.

Local Extrema detection:

To select maxima and minima of D((x,y,σ) ,each sample point is measured with eight neighbors of same image and nine neighbors in the scale above and 9 neighbors in the scale below and selected if value has maximum as maxima and minimum value as minima.



2. *keypoint localization:*

To detect stable point the method used is not only to reject keypoint with low contrast having value less than 0.03 but also uses DoG function. The DoG has strong response with edges.

$$H = \begin{bmatrix} Dxx & Dxy \\ Dxy & Dyy \end{bmatrix}$$

The Hessian matrix of f(x) is the square matrix of the second partial derivatives of f(x).

Let α be maximum Eigen values and β be the minimum Eigen values.

$$\frac{Tr(H)^2}{Det(H)^2} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2}$$
$$= \frac{(r+1)^2}{r} \quad (4)$$

The value $\frac{(r+1)^2}{r}$ is minimum when two Eigen values are equal and it increases as r ($r = \frac{\alpha}{\beta}$) and to compare the ratio of principle of curvature is below threshold r, we have to check

$$\frac{Tr(H)^2}{Det(H)^2} < \frac{(r+1)^2}{r}$$

3. *.Orientation Assignment:*

In this for each image sample L (x, y) at scale gradient magnitude m(x, y) and orientation Φ(x, y) is determined as follows:

$$M(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (5)$$

$$\Phi(x,y) = tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \quad (6)$$

4. *Keypoint description:*

The above all steps assign image location, scale, orientation and orientation to each keypoint.

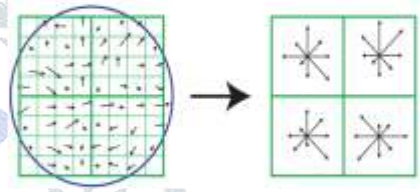


Fig.2. Image Gradient Keypoint Descriptor

As shown in fig.2 first to select level of blur in image gradient and orientation are sampled. The coordinates of the descriptor and gradient orientations are rotated in concurrence with keypoint orientation. Later, the gradients are precomputed for all levels. In fig.2 they are shown by small arrow. The Gaussian window is used to avoid sudden changes in the descriptor .It is also used to give less importance to gradients that are away from the center of the descriptor.

The David G.Lowe [2] conclusions are: a) each pixel value is multiplied by a constant will correspondingly multiply gradients by the same constant will lead to image contrast change. b) A change in brightness will affect gradient as it is determined by pixel difference. Therefore, the descriptor is not variable to affine changes in illumination but non-linear illumination changes can also occur due to camera saturation. It will cause large change in magnitudes for some gradients but have lesser effect gradient orientations. Therefore, we reduce the influence of large gradient magnitudes by thresholding the values in the unit feature vector to each by no larger than 0.2 and then renormalizing to unit length. This means that matching the magnitudes for large gradients is not important and that the distribution of orientations has greater importance. The value of 0.2 was determined experimentally using images containing differing illuminations for the same 3D objects [2].

Effect of intensity, Scale and Rotate:

1. *Intensity*

(a) (b) (a) (b)

Fig.3. a) Image after changing intensity. b) 2nd image with keypoint mapped into it.

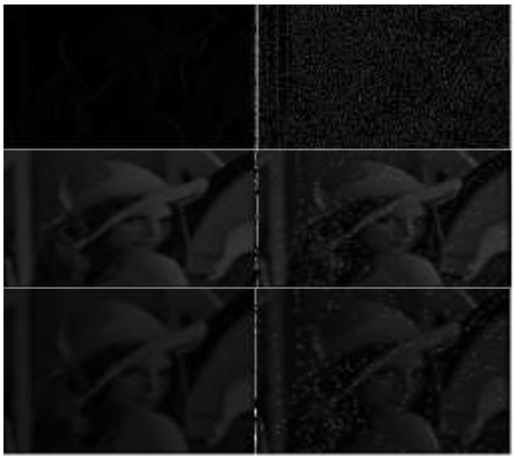
2. *Rotate*

(a) (b) (C)

Fig.4. a) Rotated image b) Straitened version of rotated image

c) 2nd image with keypoint mapped into it.

3. *Scale*



(a) (b)

Fig5.a) Image after changing scale value b) 2nd image with keypoint mapped into it.

B. Speeded-Up Robust Features:

Herbert Bay et.al [05] proposed SURF algorithm in 2006 which is used for image registration and object recognition. SURF is an algorithm to find out matching point between two images. It involves three main steps: a) interested points are selected at different location in the images. b) Neighborhood of every interested point is represented by feature vector. c) descriptor vectors are matched to different images. In SURF algorithm Hessian-Matrix approximation approach is used for interest point detection. It follows use of integral images which reduces time computation and become fast. A quick Hessian matrix is used for detection due to which speed and accuracy increased. The important thing is an integral image algorithm is used. SURF first distributes the neighborhood region of each extreme point into a number of 4×4 square sub-regions. Then, it determines a Haar wavelet output of each sub-region. The descriptor describes the distribution of intensity content within interest point neighborhood similar to scale-invariant feature transform.

Fig.6 best matching of two images by SURF algorithm

Comparison Of SIFT and SURF For Matching:-

(a) SIFT

(b)SURF

Fig.7. Comparison of SIFT and SURF algorithm for image matching

IV. CONCLUSION

The SIFT and SURF algorithm become popular and number of improvements have been made on SIFT. The modified algorithms are GSIFT, PCA-SIFT, CSIFT, SURF and ASIFT.In this paper we studied SIFT and SURF algorithms in details and studied its performance in different condition: Scale change, rotation change, illumination change. These are image local feature description algorithms based on scale-space. SIFT and CSIFT perform the well under scale and rotation change. CSIFT performs well under blur and affine changes but not perform in illumination change situation. GSIFT performs the best under blur and illumination changes. SIFT performs the best under affine change. SURF perform not enough well in different situations such as illumination changes, blur change, affine change but it is the fastest algorithm. These algorithms can be used for image registration, object recognition, and so on. In future we are using SIFT image matching algorithm to design interactive 2D-to-3D video conversion system.

V.REFERENCES

[1] Luc Vosters and Gerard de Haan,"Efficient and stable sparse-to-dense conversion for automatic 2D-to-3D conversion", IEEE transaction on circuit & system for video technology, Vol.23 No.3, March-2013, pages 373 to 386

[2] David G.Lowe, "Distinctive image features from scale-invariants keypoints", Accepted for publication in the International Journal of Computer Vision, 2004

[3] Ke, Y., Sukthankar, R. (2004), "PCA-SIFT: A more distinctive representation for local image descriptors", Computer Vision and Pattern Recognition (CVPR 2004), 27 June – 2 July 2004. IEEE, Vol. 2, 506-513.

[4] Abdel-Hakim, A.E., Farag, A.A. (2006), "CSIFT: A SIFT descriptor with color invariant characteristics", Computer Vision and Pattern Recognition (CVPR2006) , 17-22 June 2006. IEEE, Vol. 2, 1978-1983

[5] Bay, H., Tuytelaars, T., Gool, L.V. (2006), "SURF speeded up robust features", Computer Vision –ECCV 2006: 9th European Conference on Computer Vision, 7-13 May 2006. Springer, Part II, 404-417

[6] Morel, J.M., Yu, G. (2009), "ASIFT: A new frame work for fully affine invariant image comparison", SIAM Journal on Imaging Sciences, 2 (2), 438-469.

[7] Jian Wu, Zhiming Cu (2013), "A comparative study of SIFT and its variant", measurement science review, volume 13, No. 3, 2013

[8] Rafael C.Gonzalez.RichardE.Woods,Pearson edu, "Digital Image Processing",2009

Mr. R. M. Mulajkar received his BE & ME degree in Electronics Engineering in 2006 & 2011 resp. He completed PhD in Electronics Engineering at Sant Gadge Baba Amravati University, Amravati, India. His reasearch intrests include Image and Video Processing.