Key-point detection of an image using SIFT

[1]Shraddha Dewangan, [2]Anish Kumar Vishwakarma

[1]Research Scholar, [2]Assistant Professor
Department of Electronics and Telecommunication
Rungta College of Engineering of Technology, Bhilai

***Abstract*- Editing of digital images are very common these days. The process of creating fake images has been very simple with the introduction of powerful editing tools. Such tempering with digital images is known as image forgery. With the advancement of the digital image processing editing tools and software, a digital image can be easily forged. The pixel-based forgery detection aims to verify the authenticity of digital images without any prior knowledge of the original image. The detection of digital image forgery is very important because an image can be used as legal evidence, in investigations, and in many other fields.. This paper explores the effectiveness of the Scale Invariant Feature Transform (SIFT) for image matching. This method popularly used interest point detection method often applied to image matching.**

***Keywords*- image forgery, SIFT, interest point**

I. INTRODUCTION

Key-point detection is a recent terminology in computer vision that refers to the detection of interest points for succeeding processing. An interest point is a point in the image which it has clear, preferably mathematically well-founded, definition; it has a well describe position in image space, it is stable under local and global unsettled in image domain as brightness/ illumination variations, such that the interest points can be authentic computed with high degree of repeatability. After extracted interest points in two images, we are possibly able to recognize corresponding points in the two images. This process is known as image matching and it can moreover apply for point tracking, image stitching, automatic determination of popular geometry.

Image matching is a fundamental aspect of many problems in computer vision, including object or scene recognition, solving for stereo correspondence, 3D structure form multiple images, and motion tracking. Scale-invariant feature transform (SIFT) proposed by David Lowe in 2004 [1] is an algorithm for extracting interest point features from images that can be handle-down to operate reliable matching between different views of an Object or scene. The features are unaffected from rotation, image scale, and partially invariant (i.e. robust) to variation in addition of noise, 3D viewpoint, and variant in illumination. They are well defined in both the spatial and frequency domains, clutter, reducing the probability of disruption by occlusion, or noise. Large numbers of features can be extracted from typical images with well-organized algorithms. In addition, the features are highly distinctive, which allows a single feature to be precisely matched with large probability against a large database of features, delivering a basis for object and scene recognition.

II. SCALE INVARIANT FEATURE TRANSFORM

The Scale Invariant Feature Transform (SIFT) is a feature-based method for obtaining interest point features from images, that is, it not only detects interest point locations but also extracts features points that can be used to perform valid matching between different angle of an scene or object. The SIFT features are unaffected to image orientation, image scale, and provide tough matching across a considerable range of related distortion, change in 3D viewpoint, addition of noise, and change in illumination. For image matching, SIFT features are first extracted from a set of reference images and stored in a database. An image is matched by individually comparing each feature from the image to this prior database and finding candidate matching features based on Euclidean distance of their feature vectors.

A. Algorithm steps

The SIFT can be reviewed as the following four steps:

a) Scale space peak selection
b) Key-point localization
c) Orientation Assignment
d) Generation of Key-point descriptors.

Scale space peak selection: Given an input test image, SIFT features are extracted at different scales using a scale-space representation is implemented as an image pyramid. The pyramid levels are obtained by Gaussian smoothing and sub-sampling of the image resolution while feature points are selected as local extrema (minimum or maximum) in the scale-space. The first stage of key-point detection is to identify locations and scales that can be repeatable allocated under variant views of the same object. Detecting locations that are not varies to scale change of the image requires that we search for stable features across all possible variance of scale, using a continuous function of scale known as scale space. Therefore, the scale space of an image is defined as a function, $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input test image, I(x,y).

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

Where * is the convolution operation of x and y, and

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

To efficiently detect stable interest point locations in scale space, the proposed (Lowe, 1999) using scale-space local maxima and minima in the difference-of Gaussian function convolved with the image, D(x; y; σ), which can be computed from the difference of two nearby scales divided by a constant factor k:

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma)$$

In addition, the difference-of-Gaussian function provides a nearly approximation to the scale-normalized Laplacian of Gaussian, as studied by Lindeberg (1994). Lindeberg founded that the normalization of the Laplacian with the factor σ^2 is required for true scale invariance.

Keypoint localization: To detect the locations of all peaks (local maxima and minima) of $D(x,y,\sigma)$, the difference-of-Gaussian function convolved with the input image in scale space. This can be done most effectively by first octave a scale space representation that samples the function at a uniformly grid of scales and locations.

Each sample point is exmaine to see whether it is larger or smaller than all neighbours, using neighbours in both image scale and location. We have to check the eight closest neighbours in image location and nine neighbours in the scale above and below. While extremum candidates could be extracted by checking fewer neighbours, experimental results show a significant improvement in stability by selecting an extremum over this greater neighbourhood. The cost of this analysis is reasonably low due to the fact that most sample points will be removed following the first few checks.

Orientation Assignment: Once a peak candidate has been found by comparing a pixel to its neighbours, the next step is to execute a comprehensively fit to the nearby data for location, edge response, and peak magnitude. This information allows points to be unacceptable that have low contrast and therefore sensitive to noise are poorly localized along an edge. To determine this orientation, a gradient orientation histogram is computed in the neighbourhood of the key-point.

The scale of the key-point is used to pick the Gaussian smoothed image, L, with the closest scale, as all computations must be performed in a scale-invariant behaviour. For each image sample, $L_{x,y}$, the gradient magnitude, m, and orientation, $\theta(x,y)$, is pre-computed using pixel differences:

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

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

An orientation histogram is devised from the gradient orientations at all sample points within a circular window throughout the key-point. Each sample added to the histogram is loaded by its gradient magnitude and by a Gaussian-weighted circular window with an σ three times that of the scale of the key-point. The orientation histogram has 36 bins covering the 360 degree range of orientations.

Generation of Key-point descriptors: After image scale, location, and orientation have been allocated to each key-point, it is possible to enforce a two dimensional coordinate system to outline the local image region and provide undeviating with respect to these parameters. The next step is to calculate a descriptor for the local image region that is definite yet invariant to additional variations such as change in illumination and three dimensional locations.

The local image gradients are evaluated at the selected scale in the region around each interest point, and transformed into a representation that allows for local shape distortion and varies in illumination.

.One approach would be to sample the local image intensities around the key-point at the appropriate scale, and to test these using a normalized correlation measure. Although, simple correlation of image patches is highly sensitive to changes due to mis-registration of samples, such as affine or non-rigid deformations or 3D viewpoint change.

III. RESULTS

In this section, the proposed methodology providing three main kinds of the tests:
Intensity change, scale change, rotation. Tables I and II summarize the intensity change applied in the two test image. The input image is the Lena image shown at Fig. 1 Fig. 2 presents results obtained at the 1st result of intensity value changes. Fig. 3 presents results obtained Keypoint Localization. The Keypoint descriptor generates a matrix of values called a descriptor which means features possibly used for image matching. These values show that the proposed method performs satisfactorily, providing change intensity while maintaining a high rate matching percentage and time for the entire test image.

Tables III and IV summarize the scale change and rotation applied in the two test image respectively. The results shows that the proposed method performs better.

IV. CONCLUSION

A novel methodology to support image forensics investigation based on SIFT features has been proposed. Given a suspected photo, it can reliably keypoint matching if a certain region has been change intensity and, furthermore, determine the geometric transformation such as scale and rotation applied to perform such tampering. The SIFT method is invariant to image rotation, scale, and robust to change in addition of noise, 3D viewpoint, and change in illumination.

REFERENCES

[1] David G. Lowe, *Distinctive Image Features from Scale-Invariant Keypoints*, Computer Science Department (2003), 1-29

[2] S. Birchfield, *Derivation of the Kanade-Lucas-Tomasi tracking equation*, Adobe PDF document, 2 pages, January 1997.

[3] M. Brown and D. Lowe, *Recognising panoramas*, In Proceedings of IEEE International Conference on Computer Vision (2003), no. 4, 1.

[4] P. J. Burt and E. H. Adelson, *The laplacian pyramid as a compact image code*, IEEE Transactions on Communications COM-31 (1983), no. 4, 532–540.

[5] C. Harris and M. J. Stephens, *A combined corner and edge detector*, In Alvey Vision Conference (1988), 147– 152.

[6] H. Ling and D. Jacobs, *Deformation invariant image matching*, IEEE International Conference on Computer Vision 2 (2005), 1466–1473

[7] D. G. Lowe, *Object recognition from local scale invariant features*, In International Conference on Computer Vision (1999), 1150–1157.

[8] B. D. Lucas and T. Kanade, *An iterative image registration technique with an application to stereo vision*, International Joint Conference on Artificial Intelligence (1981), 674–679.

[9] J. Matas, P. Bilek, and O. Chum, *Rotational invariants for wide-baseline stereo*, Research Report of CMP (2003), 1.

[10] H. P. Moravec, *Visual mapping by a robot rover*, International Joint Conference on Artificial Intelligence (1979), 598–600.

[11] J. Shi and C. Tomasi, *Good features to track*, IEEE Conference on Computer Vision and Pattern Recognition (1994), 1–10.

[12] C. Tomasi and T. Kanada, *Detection and tracking of point features*, Carnegie Mellon University Technical Report CMU-CS-91-132 (1991), 1–10.

[13] T. Tuytelaars and L. Van Gool, *Wide baseline stereo based on local, affinely invariant regions*, British Machine Vision Conference (2000), 412–422.

TABLE I. INTENSITY VALUE INCREASE

SN	Intensity value	Test image-1		Test image-2	
		Matching percentage	*Processing time*	*Matching percentage*	*Processing time*
1	2	100	0.0232	100	0.0073
2	3	100	0.0114	100	0.0072
3	4	100	0.0116	100	0.0080

TABLE II. INTENSITY VALUE DECREASE

SN	Intensity value	Test image-1		Test image-2	
		Matching percentage	*Processing time*	*Matching percentage*	*Processing time*
1	2	100	0.0114	100	0.0037
2	3	100	0.0218	100	0.0053
3	4	100	0.0613	100	0.0057

TABLE III. SCALE CHANGE

SN	Scale (sigma value)	Test image-1		Test image-2	
		Matching percentage	*Processing time*	*Matching percentage*	*Processing time*
1	1	51.62	0.004	28.76	0.01
2	1.3	73.87	0.0015	70.95	0.007
3	1.5	94.10	0.0009	87.53	0.004
4	1.6	100	0.0009	100	0.005

TABLE IV. ROTATION

SN	Rotate (angle)	Test image-1		Test image-2	
		Matching percentage	*Processing time*	*Matching percentage*	*Processing time*
1	30	65.11	0.003	41.23	0.009
2	45	73.17	0.002	64.09	0.012
3	90	100	0.001	100	0.005
4	135	73.17	0.001	64.09	0.010
5	180	100	0.0009	100	0.005

fig1: test image

fig 2: image after changing intesity

Fig 3: keypoints of image