IKDSIFT: An Improved Keypoint Detection Algorithm Based-on SIFT Approach for Non-uniform Illumination

Krittachai Boonsivanon\textsuperscript{a}, Anupap Meesomboon\textsuperscript{a,*}

\textsuperscript{a}Department of Electrical Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen, 40002, Thailand

Abstract

In this paper, we propose an improved keypoint detection algorithm of object-based recognition for non-uniform illumination, called IKDSIFT, which is implemented using the SIFT approach, morphological operations, Top-Hat filtering and various techniques in pre-processing procedures. The number of keypoint rate of data sets was compared. Data sets consist of three hundred 150x150 images and thirty 851x566 images with different uniform and non-uniform illumination. The experimental results show that the number of keypoint detection is reciprocal to peak selection thresholds. The best algorithm is the proposed IKDSIFT, followed by the SIFT. The ASIFT performs the worst. Additionally, the SIFT and ASIFT can detect some peak selection thresholds while the IKDSIFT can detect all ranges of the peak and obtains the best result comparing to other ones. Hence, the proposed algorithm looks promising to be used for recognizing under non-uniform illumination.

Keywords: Object-based recognition; Keypoint detection; Non-uniform illumination; SIFT approach.

1. Introduction and Related work

Object class detection has become one of the most focused areas and one of the fundamental challenges in

* Corresponding author. Tel.: +66-43-347-031; fax: +66-43-347-032.
E-mail address: anupap@kku.ac.th.
computer vision in the new century. They have been utilised in many real-world applications, for instance, NASA Mars Rover images, image reconstruction, robot localisation, video data mining, building panoramas stitching\(^4\), etc. Because of many factors such as the different absorption, reflection properties, and background noise caused by uneven illumination, many approaches have been proposed to overcome these issues\(^2\). So, object-based recognition leads to many challenging problems. Elementary characteristics that are hopefully invariant over different appearances are different locality, distinction, quantity and invariance, especially non-uniform illumination conditions. By 2004, David G. Lowe proposed a local feature description approach known as Scale-invariance Feature Transformation (SIFT)\(^3\). However, the SIFT approach performs the best under scale and rotation changes, but not illumination change\(^4\) since normalising the vector is caused by the illumination changes. Hence, these problems are investigated in this paper. The SIFT is based on a local feature description approach\(^5\) that is known for its invariance under rotation, translation, scale changes, blur changes, affine transformation, illumination changes, and other transformations. The procedure of SIFT consists mainly of four steps: 1) scale-space extrema detection, 2) keypoint localisation, 3) orientation assignment, and 4) keypoint descriptor. Firstly, the SIFT uses a Difference of Gaussian (DoG) function, Eq. (1), to do convolution on the image. We obtain different scale images by changing \(\sigma\).

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

Then, the images which are close in the similar resolution are subtracted to get a DoG pyramid. In other words, subtracting an image from its more-blurred neighbour image gives the DoG. Finally, points that are maximum or minimum in their 3x3x3 = 27 neighbourhood—9 pixels in the less-blurred image, 9 pixels in its own image and 9 pixels in the more-blurred image are marked. The DoG function is a kind of an improvement of a Gauss-Laplace algorithm (see Eq. (2))

\[
D(x,y) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \tag{2}
\]

where \(I(x,y)\) denotes an input image, and \(k\) denotes a scale coefficient of an adjacent scale-space factor. Secondly, points that have poorly localised along an edge are rejected. The interpolation to locate the keypoint accurately in scale and space is then deployed. Thirdly, \(m(x,y)\) assigns a direction to each keypoint based on local image gradients and \(\theta(x,y)\) creates a 36-bin orientation histogram and looks for peaks in the histogram (Eq. (3)). It is possible for a keypoint to be assigned multiple orientations.

\[
m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{3}
\]

\[
\theta(x,y) = \tan^{-1}\left(\left((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))\right)\right)
\]

Finally, each keypoint is summarised in a way which allows it to be compared with other keypoints, while retaining its various robustness properties. Then we calculate 16 separate orientation histograms in a 4x4 neighbourhood around each keypoint. The histograms are calculated with respect to the keypoint scale and orientation which have been distinct in previous steps. Each histogram has 8 orientation bins. The contents of all of the histograms are concatenated to form a 128-element (16x8) vector. This vector is called the keypoint descriptor. Normalising the vector makes it more robust to illumination changes. Now, we will briefly discuss the related work on a local scale invariant features. Firstly, proposed Top-Hat transformation\(^5\), which consists of steps as follows: 1) Let \(f(x)\) and \(b(x)\) be two discrete functions defined on two-dimensional discrete space \(F\) and \(B\), respectively, 2) The opening (\(\circ\)) and closing (\(\bullet\)) morphological operations are applied to \(f(x)\) and \(b(x)\), 3) The Top-Hat operator is divided into opening and closing Top-Hat operators (\(OTH_{fb}\) and \(CTH_{fb}\), respectively) defined by

\[
OTH_{fb}(x) = (f \circ b)(x) \tag{4}
\]

\[
CTH_{fb}(x) = (f \bullet b \circ f)(x) \tag{5}
\]
Like SIFT, Affine-SIFT, called ASIFT, uses grayscale information of images and deploys affine transformation parameters to correct images and intends to resist strong affine changes. ASIFT simulates the rotation of camera optical axis. It naturalises an image affine transformation model resulting from the changes of viewpoint. Wu et al. proposed systematic analysis on the major members of the SIFT and its variants, including PCA-SIFT, GSIFT, CSIFT, SURF and ASIFT according to performance evaluation and comparisons under different situations: scale, rotation, blur, affine and illumination. The comparative results showed that the SIFT performs the best under scale and rotation changes while ASIFT performs the best under affine transformation. However, both do not perform well under the illumination change. Therefore, SIFT and ASIFT approaches are selected and investigated in this paper since it is also a necessary precondition to solve several problems. Wang et al. proposed a novel method in machine vision system based on homomorphic filtering, Top-Hat transformation and watershed algorithm. Experimental results showed that the method is simple and practical; it can segment the targets from uneven illumination particle images. Finally, Gupta et al. proposed and evaluated an accurate and simple algorithm to identify the microscopic particles present in an image and compute area based statistics of each and every particle using morphological operation and image enhancement features that due to non-uniform background illumination. However, this algorithm based on ‘region of interest’—ROI could not differentiate between some of the particles and their background or neighbouring pixels.

2. The proposed algorithm

The IKDSIFT algorithm that is implemented using the SIFT approach consists of mainly following steps—scale-space extrema detection and keypoint localisation, morphological operations, Top-Hat filtering and various techniques is organized as follows: 1) an image is acquired from the camera and converted to gray scale, 2) the background is estimated by the morphological opening operation and Top-Hat filtering using the similar structuring element for both operations so that the background is brighter in the centre of the image, 3) background image is subtracted from the original image so that step 2) and step 3) together could be followed by Top-Hat filtering in next step, 4) Top-Hat filtering, 5) thresholding so as to remove the background noise, 6) identifying objects in the image, we then create a label matrix and display it as a pseudo-color indexed image (color map matrix). Finally, keypoints are described.

3. Experimental setup and results

We compare the number of keypoint detection rate of data sets provided by the Purdue RVL SPEC-DB and the PHOS color image database. Data sets consist of three hundred 150x150 images and thirty 851x566 images. The data sets were used to test uniform and non-uniform illumination invariance. The prepared parameters were set the peak selection thresholds which were of the values 0, 3.5, 5, 7.5 and 10. All experiments were performed on a 2.13 GHz Intel(R) Core i5 CPU and 4GB RAM under Windows7. The IKDSIFT and its variants were implemented using C API and MATLAB software which was complied with VLFeat library. The comparison of the number of keypoint detection is shown in Fig.1., we can see that the performance of IKDSIFT is the best, followed by SIFT and ASIFT, respectively. The results under non-uniform illumination situations show that the best algorithm is the IKDSIFT. The maximum number of keypoint detection is 1,551 keypoints and minimum of 161 keypoints obtained by IKDSIFT, followed by the SIFT. The ASIFT performs the worst. Additionally, the SIFT and the ASIFT can detect peak selection thresholds between 0 to 5 and 0 to 7.5, respectively. The IKDSIFT can detect all ranges of the peak (0-10) and obtain the best result comparing to others.

4. Summary

The experimental results show that the number of keypoint detection is reciprocal to the peak selection thresholds. The best algorithm is the proposed IKDSIFT, followed by the SIFT. The ASIFT performs the worst. Additionally, the SIFT and the ASIFT can detect some peak selection thresholds while the IKDSIFT can detect all
ranges of the peak and obtain the best result comparing to other ones. Therefore, IKDSIFT looks promising to be used for recognizing under non-uniform illumination.

![A comparison of keypoint detection](image)

**Fig. 1.** The comparison of number of keypoint detection under non-uniform illumination (peak threshold = 5), (a) IKDSIFT (326 keypoints); (b) SIFT (57 keypoints); (c) ASIFT (22 keypoints).

**Acknowledgements**

This research is financially supported by Khon Kaen University under the Incubation Researcher Project.

**References**