1. **STFT algorithm for keypoint detection**

The keypoint detection function is implemented using SIFT algorithm provided by OpenCV. The SIFT(Scala-Invariant Feature Transformation) algotrithm is a technique introduced by David Lowe in 1999 and is widely used for extracting features of images in the field of computer vision. The SIFT algorithm consists four major steps: Scale-space extrema detedtion; Keypoint localization; Orientation assignment; and Keypoint descripter.

* 1. **Scale-space extrema detection**

The fisrt step in SIFT algorithm is convoling image with a Guassian kernel to construct an image pyramid, which ensures that each SIFT feature has corresponding relationship at different scale[[[1]](#endnote-0)]. The method is Difference od Guassians(DoG). Given the original image *I(x,y)*, DoG is defined as follows:

where *L(x,y,kσ)* is obtained by convolution the original image *I(x,y)* and the Guassian kernel *G(x,y,kσ)*.

* 1. **Keypoint localization**

Due to the discreteness of scale sapce and pixels, the Duassian difference pyramid is also discrete. Some potential keypoints are therefore unstable. Talyor series expansion is used to determine the position of keypoints. The Taylor expansion is given by:

where *X=(x,y,σ)* and *f* is Difference of Guassians. The derivative is:

Let it be 0 and the position of extreme points can be obtained.

* 1. **Orientation assigment**

The algorithm adopts gradient histogram method. It calculates the gradient of pixels firstly within a cell. After completing the gradient calculation of the keypoints, it uses the histogram to count the gradient and direction of the pixels in the neighborhood. The gradient histogram divides the direction range from 0 to 360 degrees into 36 bins, each of which represents a range of 10 degrees.

* 1. **Keypoint descripter**

In this step, the patch is divided into 4×4 blocks, each of which is 16×16 pixels. Each block has an 8-bin histogram to store the gradient orientations. Histogram increases by weighted magnitude for each gradient.

The keypoint detection function converts the image to gray and then uses it as input of SIFT function. It returns the coordinate of kepoints. One input and output images are shown as follows:



Figure 1. Original image

 Figure 2. Output of SIFT

1. **Feature bucketing**

Feature bucketing is a method that ensures feature points are well distributed in image and reduces computation complexity of algorithm. It divides the image into several cells, which are called buckets. For each bucket, it selects 10 points with strongest response and then return the coordinte of keypoint after bucketing. Response indiactes how good the keypoints are. The strongest response corresponds the best keypoint. The followwing is the ouput of feature bucketing method:

 Figure 3. Ouput of bucketing

Compared with Figure 2, the number of keypoinys in Figure 3 is reduced.

1. **Feature tracking**

Feature tracking is a problem to track keypoints in an image sequence, which can be used for motion analysis, object tracking, etc. One of the important method is called Lucas-Kanade method, which is based on Optical Flow.

* 1. **Optical Flow**

Suppose the pixel *(x,y)* at T moves to *(x+dx,y+dy)* after time dt. According to britness consistancy:

where *I(x,y,t)* is britness of pixel *(x,y)* at time t. Applying Taylor expansion:

then we have:

and the Optical flow equation is:

* 1. **Lucas-Kanade method**

The main principle of Lucas-Kanade optical flow estimation is to assume constant brightness to find the velocity vector between two consecutive frames (t and t+1)[[[2]](#endnote-1)] by the following equation:

Python library OpenCV provide a function *calcOpticalFlowPyrLK()*, which implements Lucas-Kanade method. The function uses the coordinate of keypoints at time t that returned by SIFT function in the last step to calculate the position of the same keypoints at time t+1. It returns the value of status for each input point. If the corresponding keypoint is found, the value of status will be 1.

After that, the feature tracking function will return the coordianate of corresponding points with status equal to 1. The data obtained will be used to generate 3D point cloud.

1. [] Abdullah-Al N, Kong Y . Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network[J]. Information, 2018, 9(1):19-19. [↑](#endnote-ref-0)
2. [] O. Haggui, C. Tadonki, F. Sayadi and B. Ouni, "Evaluation of an OPENMP Parallelization of Lucas-Kanade on a NUMA-Manycore," 2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD), Lyon, France, 2018, pp. 436-441, doi: 10.1109/CAHPC.2018.8645936. [↑](#endnote-ref-1)