

# ML

YUCHIH YU, ycy247

May 14, 2014

1. SIFT stands for Scale-invariant Feature Transform, it's an algorithm that extracts Keypoints and computes its descriptors from an image and Keypoints are just some points in an image that we think can represent the image. The characteristic of this algorithm is that it can find Keypoints in different scales.

Here's how SIFT works. First of all, we know that, in order to find Keypoints in different scales we can use LoG(Laplacian of Gaussian) by changing the value of sigma. But that would be costly, so there's another way to approximate the result of LoG and it's called Dog(Difference of Gaussians). And the location of Keypoints are found by finding the local extrema over scale and space. After the locations are found, some noise Keypoints(such as edges and low contrast points) are eliminated. And then the direction of each Keypoint can be determined by picking the highest bins of 36 bins histogram based on image gradient orientation. Now, for each Keypoint, we can create a 128 bin vector to represent it by the 4X4 neighboring cell and 8 bins orientation histogram. Now we have features so we can match features by using nearest neighbors(only those that the closest neighbor is much closer than the second neighbor).

In the data set that we are experimenting, same objects in the image may have different sizes and orientations. So we think SIFT is a good base line for our experiment. For the parameter, the author of SIFT paper have done many experiments on them and suggested the best ones to use, so we decided to stick with it.