

## Concept of K Means

The K means machine learning algorithm is used to group a definite number of data points into a given number of clusters based on mere relative positioning of the observations from each other.

Basically the purpose of this algorithm is to find

EQUATION:

$$\arg \min_S \sum_{count=1}^k \sum_{x \in S_i} ||x - \mu_{count}||$$

Where

“x” is a set of dataset points

“k” is the total number of clusters formed by the algorithm.

“S” is the set of clusters formed.

“ $\mu_i$ ” is the mean of points in “ $S_i$ ”.

Such that the variance of a particular cluster is minimum (Minimum “Within cluster sum of squares” [WCSS])

The use of K means in our algorithm plays 2 very important roles in our algorithm.

1. Increase the probability of precisely locating the object
2. Enable us to determine the boundaries of the object.

1. Increase the probability of precisely locating the object

Once we obtained the MATCHED FEATURE SURF DESCRIPTORS of object scene image pair from the algorithm we run K means algorithm onto the LOCATION of these points.

As the density of the matched feature SURF descriptors shall be higher at the location of object we choose the cluster with highest density to extract the information of the location of the object.

2. Enable us to determine the boundaries of the object.

The minimum Eigenvalue feature algorithm outputs more features than the SURF feature as it is based on corner detection than blob detection.

The location of these points are fed into the K means clustering algorithm thus outputting the clusters of each individual object separately from the scene.

Thus the use of K means algorithm enables us to precisely map the boundaries of the object better than the combination of SURF and K Means.