

Scale-invariant feature transform (or SIFT) is an algorithm in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999. SIFT is a method to detect distinct, invariant image feature points, which easily can be matched between images to perform tasks such as object detection and recognition, or to compute geometrical transformation between images.

SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors.

Lowe's method for image feature generation transforms an image into a large collection of feature vectors, each of which is invariant to image translation, scaling, and rotation, partially invariant to illumination changes and robust to local geometric distortion.

The scale invariant features are efficiently identified by using a staged filter approach. The first stage identifies key locations in scale space by looking for locations that are maxima or minima of a difference of-Gaussian function. Each point is used to generate a feature vector that describes the local image region sampled relative to its scale-space co-ordinate frame.

The SIFT keys derived from an image are used in a nearest-neighbour approach to indexing to identify candidate object models. Collections of keys that agree on a potential model pose are first identified through a Hough transform hash table, and then through a least-squares fit to a final estimate

of model parameters. When at least 3 keys agree on the model parameters with low residual, there is strong evidence for the presence of the object.

KeyPoints Matching using FLANN:

After successfully finding SIFT keypoints in an image, we can use these keypoints to match the keypoints from other image.

1. Select ROI in the image. Slice the image and extract the ROI.
2. Find SIFT Keypoints in ROI image.
3. Find SIFT Keypoints in the image.
4. Match Keypoints using FLANN.
5. Predict new bounding box/ ROI and repeat.

FLANN:

Fast Library for Approximate Nearest Neighbors is a library that contains a collection of algorithms optimized for fast nearest neighbor search in large datasets and for high dimensional features.

Other Important Links:

1. SIFT – <http://www.cs.ubc.ca/~lowe/keypoints/>
2. SIFT code – <http://blogs.oregonstate.edu/hess/code/sift/>
3. FLANN – <http://mloss.org/software/view/143/>