Feature Descriptors

There is a large list of ways to compute the descriptors, and represent them.

- HOG : Histogram of oriented gradients.
- SIFT : Scale invariant feature transform.
- SURF : Speeded up robust features.
- BRIEF : Binary robust independent elementary features.
- ORB : Oriented fast and rotated brief.
- Many more...

- SIFT.
  - scale invariant: descriptor that is invariant to translation, rotation and scale.
  - partially invariant to affine transformation, illumination changes and 3D projection.
  - suitable for detecting visual landmarks from different orientations, and distances, with different illuminations.

  - SIFT Feature <p, s, r, f>

  p - pixel location.

  s - scale. Extrema in scale space from DoG.

  r - orientation. Peak from a histogram of gradients in a local region.

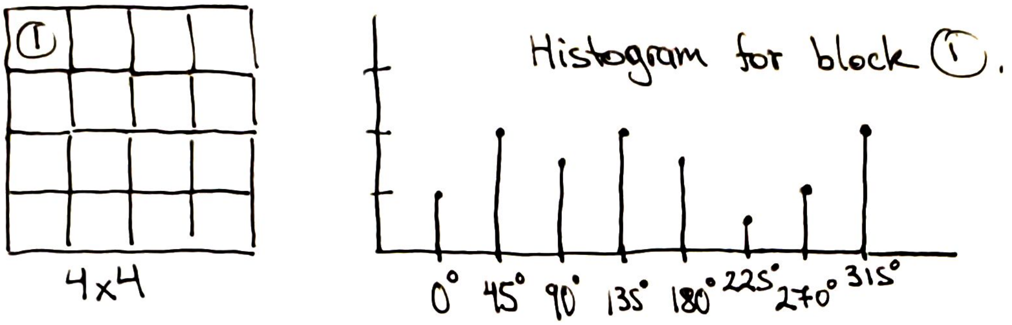
  f - 128 dim, descriptor, computed from local image gradient.

  - As you can see p, s and r are view point dependent. And f is mainly independent of a view point.

  - SIFT is expansive to compute, and its patented.
  - SIFT is a golden standart, SURF is a faster variant.

- SIFT Descriptor

8 orientations × (4×4) historgram array = 128 dimensions.



Hence, we have (4×4) = 16 histograms with 8 entries each.

For a total of 128 values.

- Lowe's Ratio Test.

A 3 step test to eliminate ambiguous matches for a query feature q.

1. Find the closest 2 descriptors P_1 and P_2 based on euclidian distance d.
2. Test the distance to best match against a threshold, $d(q, P_1) < T$.
3. Accept a match only if the best match is substantially better than second.

$$\frac{d(q, P_1)}{d(q, P_2)} < \frac{1}{2}$$

, $\frac{1}{2}$ is the usual distance ration threshold.

Meaning our best match needs to be twice as close as our second best match.

Binary Descriptors.

- Select a area around a pixel, then select a number of pairs between the pixels in the area.

- Compute comparison between each intensity in pair.

$$b = \left\{ \begin{matrix} 1 & \text{if } I(s_2) > I(s_1) \\ 0 & \text{else} \end{matrix} \right\}$$

Concatenate all b's to a bit string.

If (s_1, s_2)

Pairs $(5,1)(5,9)(4,6)(8,2)(3,7)$

$b = \quad 0 \quad 0 \quad 0 \quad 1 \quad 1$

Example

0	255	100
100	100	0
255	0	50

Intensities

1	2	3
4	5	6
7	8	9

index

Final descriptor is 00011.
B = 00011.

This results in a very concise descriptor that is computed using pairs that are chosen based on certain strategy.

- Pros
  1. compact descriptor.
  2. fast to compute.
  3. trivial and fast to compare. Can be done using Hamming distance.

$$d_H(B_1, B_2) = \text{sum}(\text{xor}(B_1, B_2))$$

- Stragedy must be fixed, same pairs, the order of chosen pairs must be mainted from one image to another. Most binary descriptors mainly differ by stragedy.

- Examples of binary descriptors is BRIEF, and ORB.