Page No.
Date: / /

Visual Features: Descriptors (SIFT, BRIEF, and ORB)

SIFT Descriptor

⇒ Image content is transformed into features that are invariant to:
- → Image translation
- → Rotation
- → Scale

⇒ They are partially invariant to:
- → Illumination change
- → affine transformations & 3D projections.

★ SIFT Features

⇒ A SIFT feature is given by a vector computed at a local extreme point in the scale space.

$\langle P, S, r, f \rangle$

- P: Pixel location of the keypoint in the image
- S: Extrema in the scale-space from DoG
- r: Orientation: peak image gradient direction around keypoint
  - → ① Compute image gradient in a local region
  - → ② Build histogram
  - → ③ Select peak as the keypoint orientations
- f: 128 dim descriptor generated from local image gradients

→ Independent

→ View-point dependent

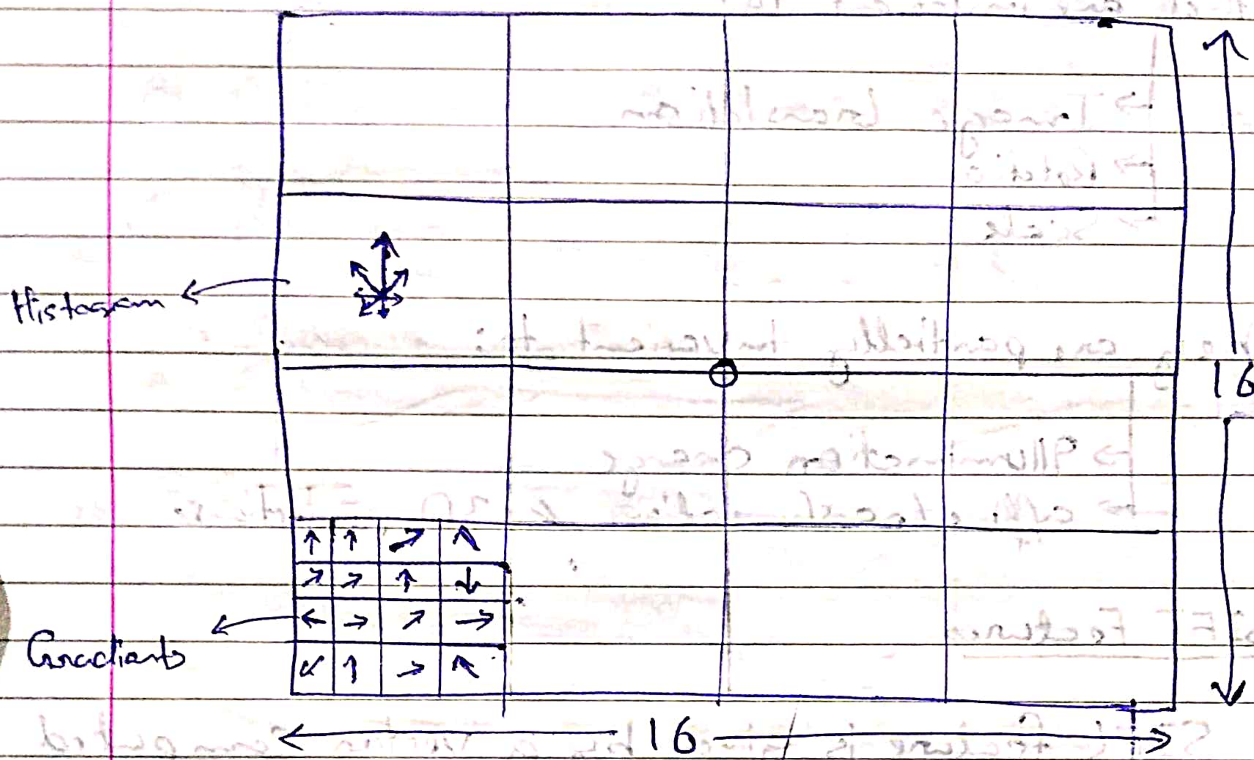
20:00

Page No.

Date: / /

★ SIFT Descriptor In Sum

⇒ Compute image gradients in local 16×16 area at the selected scale.



⇒ Create an array of orientation histograms

↳ 8 orientation × (4×4) histogram array = 128 dimensions (yields best results)

★ Lowes Ratio Test

⇒ 3 Step test to eliminate ambiguous matches for a query feature q.

1. Find closest two descriptors to q, called P_1 and P_2 based on the Euclidian distance d.

2. Test if distance to best match is smaller than a threshold: $d(q, P_1) < T$.

3. Accept match only if the best match is substantially better than second.

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

Binary Descriptors { Computing descriptor fast }

⇒ Complex features such as SIFT works well and is a gold standard but it is expensive to compute.

★ Key idea of binary descriptor

⇒ Select a patch around a Keypoint.

⇒ Select a set of pixel pairs in that patch

⇒ For each pair, compare the intensities.

$$b = \begin{cases} 1 & \text{if } I(S_1) < I(S_2) \\ 0 & \text{otherwise} \end{cases}$$

⇒ Concatenate all b's to a bit string.

★ Key advantages of Binary Descriptors

① Compact description

↳ The number of pairs gives the length in bits.

② Fast to compute
↳ Simply intensity value comparisons.

③ Trivial and fast to compare
↳ Hamming distance ($d_{Hamming}(B_1, B_2)$)

⇒ Different binary descriptors differ mainly by the strategy of selecting the pairs

BRIEF {Binary robust independent elementary features}

⇒ First binary image descriptor.

⇒ Proposed in 2010

⇒ 256 bit descriptor

⇒ Operations performed on a smoothed image to deal with noise

★ BRIEF Sampling Pairs

① GI: Uniform random sampling

② GII: Gaussian sampling

③ GIII: S_1: Gaussian, S_2: Gaussian centered around S_1

④ G IV: Discrete location from a course polar grid.

⑤ G V: S_1: (0,0): S_2 are all location from a course polar grid.

⇒ BRIEF fails when camera is rotated!

ORB {Oriented FAST Rotated BRIEF}

⇒ An extension to BRIEF that
- ↳ Adds rotation compensation
- ↳ Learns the optimal sampling pairs.

★ <u>ORB: Rotation Compensation</u>

⇒ Estimates the center of mass and the main orientation of the area/patch.

⇒ Image moment:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y)$$

⇒ Center of mass $C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$

⇒ Orientation $\theta = atan2(m_{01}, m_{10})$

Page No.
Date: / /

⇒ Given CoM and orientation C, θ, we can rotate the coordinate of all pairs by θ around C:

$$S' = T(C, \theta) S$$

⇒ Use the transformed pixel coordinates for performing the test.

⇒ Invariance to rotation in the plane.

★ ORB: Learning Sampling Pairs

⇒ Pairs should be / have:

① Uncorrelated

↳ So that each new pair adds new information to the descriptor.

② high variance

↳ It makes a feature more discriminative.

⇒ ORB defines a strategy for selecting 256 pairs optimizing for both properties using a training database.