Feature Detection & Description

PABLO MALDONADO

What are features?

➔ Something that changes a lot with respect to its (close) neighbors.

- Corners.
- Edges.
- Color changes.
- Curvature.
- …

- **Two tasks:** detect features, once we have them, describe the image in these terms.
  - **Example:** “On the left of this edge, everything is blue, on the right, is gray”.

Detection vs Description

KEYPOINT/FEATURE DETECTION

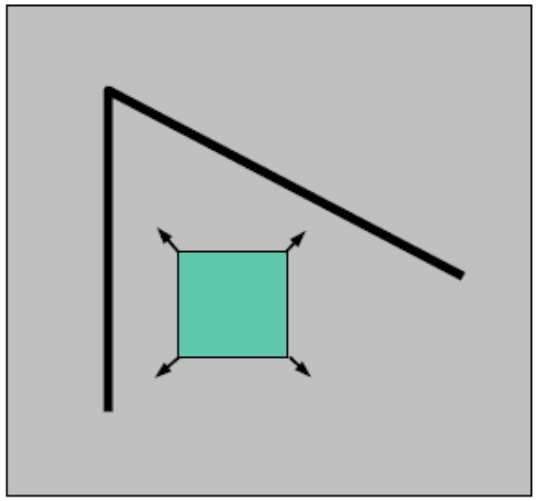
A keypoint (interest point) is a point within the point cloud that has the following characteristics:

1. it has a clear, preferably mathematically well-founded, definition,
2. it has a well-defined position in image space,
3. the local image structure around the interest point is rich in terms of local information contents.

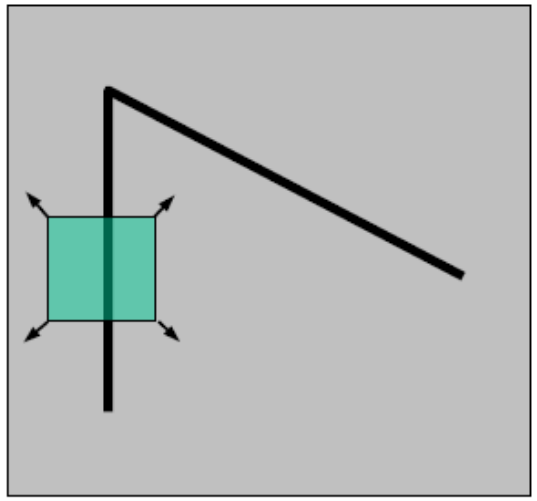
FEATURE DESCRIPTORS

After detecting keypoints we go on to compute a descriptor for every one of them. "**A local descriptor is a compact representation of a point's local neighborhood.** In contrast to global descriptors describing a complete object or point cloud, local descriptors try to resemble shape and appearance only in a local neighborhood around a point and thus are very suitable for representing it in terms of matching." (Dirk Holz et al.).

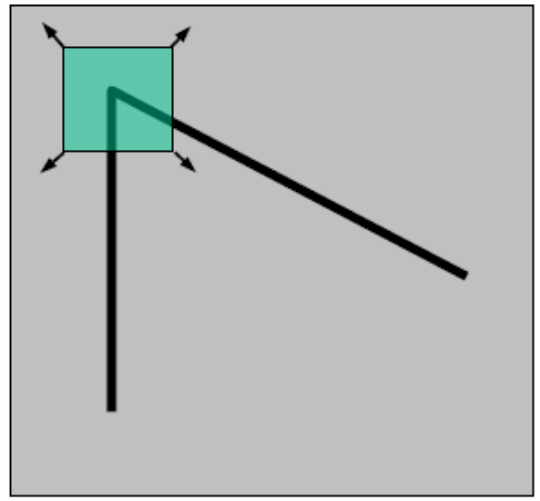
Interesting points in an image



"flat" region:
no change in all directions



"edge":
no change along the edge direction



"corner":
significant change in all directions, i.e., even the minimum change is large

Corner Detection

Harris & Stephens, 1988

- Corner = Local maximum of the difference of intensity

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} \underbrace{[I(x+u,y+v)}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

- Linearizing, we get:

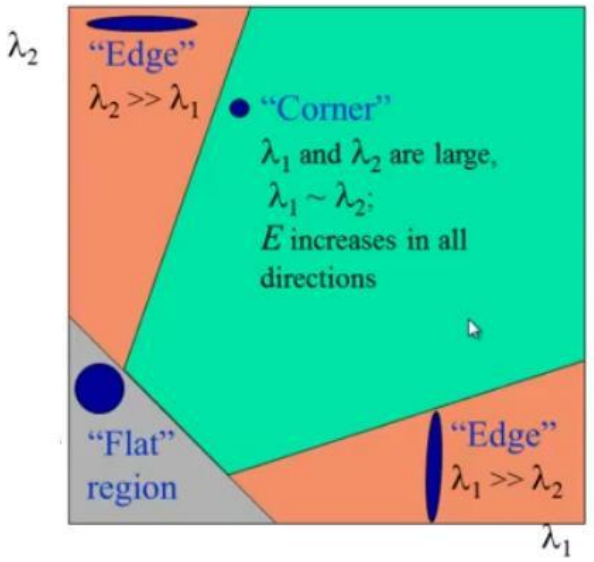
$$E(u,v) \approx \begin{bmatrix} u & v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \qquad M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

- A score is calculated (depending on a parameter):

$$R = det(M) - k(trace(M))^2$$

Harris & Stephens, 1988

- When $|R|$ is small, which happens when λ_1 and λ_2 are small, the region is flat.
- When $R < 0$, which happens when $\lambda_1 >> \lambda_2$ or vice versa, the region is edge.
- When R is large, which happens when λ_1 and λ_2 are large and $\lambda_1 \sim \lambda_2$, the region is a corner.

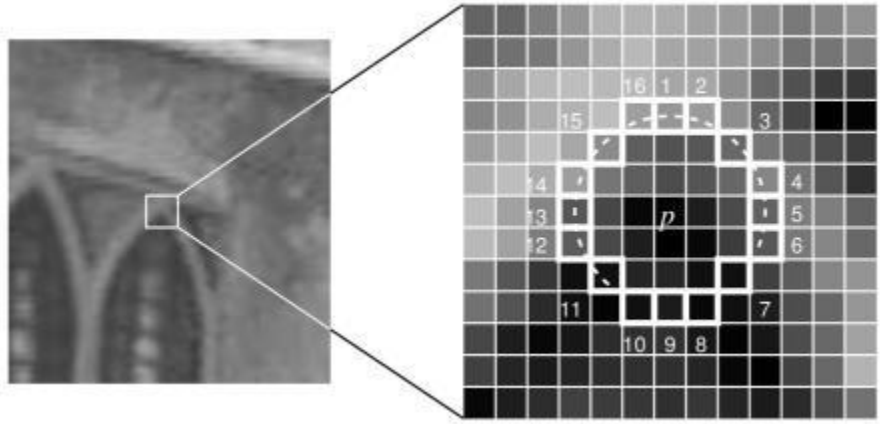


FAST

Choose a candidate pixel. Let ***I*** denote its intensity and ***t*** a threshold (hyperparameter).

The pixel is a corner if there are at least ***n*** neighbors that are all darker or all lighter than the candidate pixel, minus/plus the threshold.

For speed, check for the intensity of points 1, 9, 5, 13. Three of those should be darker or lighter.



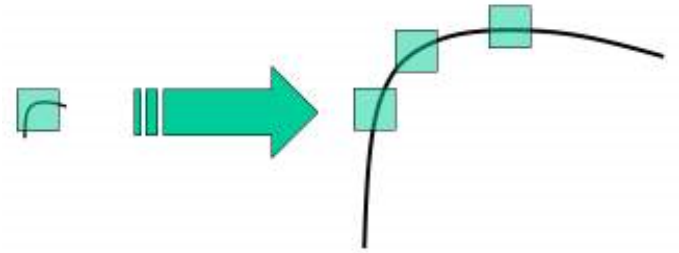
Feature descriptors

SIFT

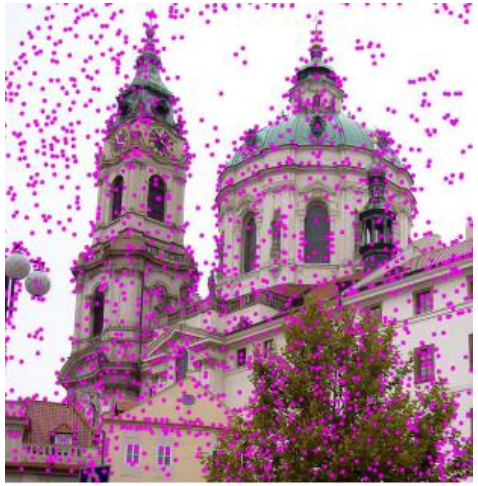
- Corners are rotation-invariant, which is useful, for instance, for motion tracking.

- But they are not scale invariant!

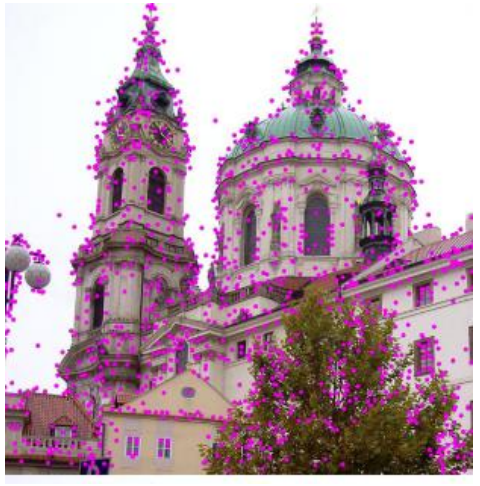
- **Could we get better quality, scale invariant features?**



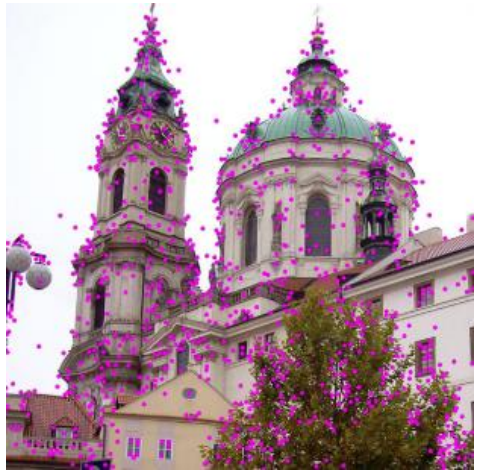
SIFT



Find local extrema across different scales (think of image pyramids).



Remove low contrast points.



Remove points on edges.

Properties of SIFT

- Robust matching technique: can handle viewpoint changes (up to 60 degrees out of rotation plane).
- Can handle significant changes in illumination (sometimes between day/night).
- Runs in real time!
- Successor: SURF (some computational improvements).

License

- Both SIFT and SURF are available in OpenCV-contrib, not in OpenCV. This is somehow an attempt to make users aware that both algorithms are patented.

- For commercial use, you should contact patent holders!

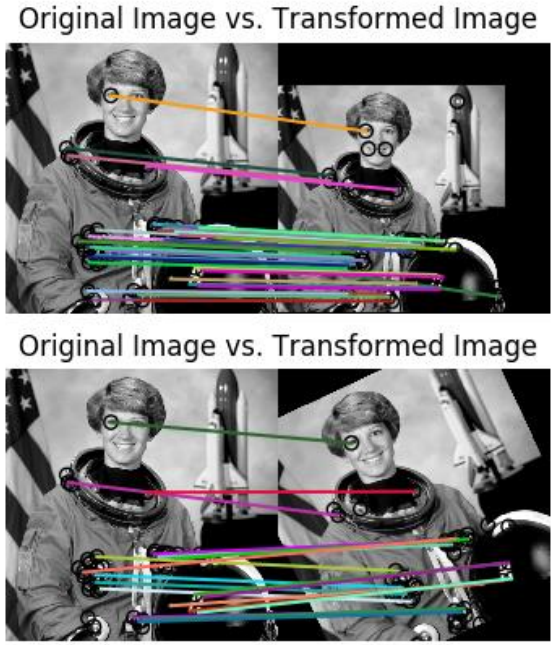
- Other (free!) alternatives: **ORB**

BRIEF

- If SURF/SIFT are great, why not use them for everything?

- Two issues:
  - Patent.
  - Computations made on the full histogram of the local patches ➔ costly computation.

- Alternative: encode into binary vectors using only comparison between intensities.



STAR / CENSURE Feature detector

Original Image

Transformed Image

ORB

- ORB uses an oriented FAST detection method and a variant of BRIEF descriptors that is robust to rotation.

- It is preferred over BRIEF for real-time applications.