# Local Image Descriptor

## Course: Computer Vision

E. Francisco Roman-Rangel.
francisco.roman@itam.mx

Digital Systems Department.
Instituto Tecnológico Autónomo de México, ITAM.

September 21st , 2020.

# Outline

ORB descriptor

Scale-Invariant Feature Transform (SIFT)

Shape Context

# Intro

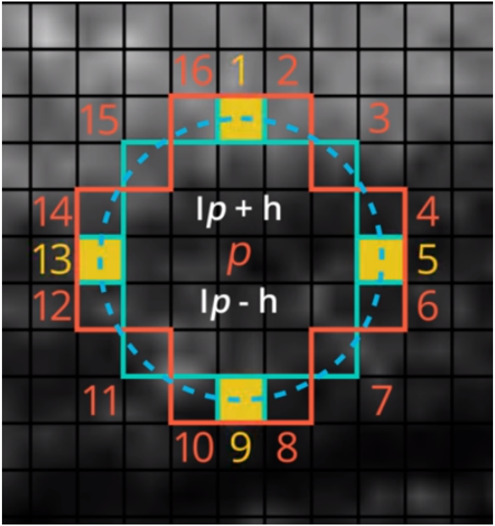
Remember the need for local image descriptors (vs using only PoI).

"Oriented FAST and rotated BRIEF".
By Rublee et al. in 2011.

Pipeline:

1. PoI detection with FAST.
2. Local orientation estimation by moments.
3. Local description by BRIEF (aligned).

# FAST



# Local orientation

First, the moments of a patch are defined as:

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

ORB descriptor-Patch's moment's definition

With these moments we can find the centroid, the "center of mass" of the patch as:

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

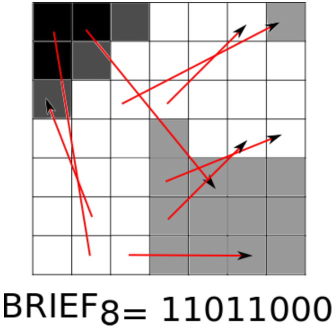
ORB descriptor — Center of the mass of the patch

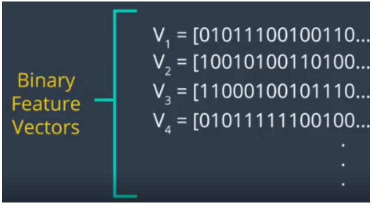
We can construct a vector from the corner's center O to the centroid -OC. The orientation of the patch is then given by:

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

ORB descriptor — Orientation of the patch

# ORB




$$\text{Where } \tau(p;x,y) \text{ is defined as:}$$

$$\tau(p;x,y) = \begin{cases} 1 & : p(x) < p(y) \\ 0 & : p(x) \geq p(y) \end{cases}$$

$$p(x) \text{ is the intensity value at pixel } x.$$

# Outline

ORB descriptor

Scale-Invariant Feature Transform (SIFT)

Shape Context

# Overview

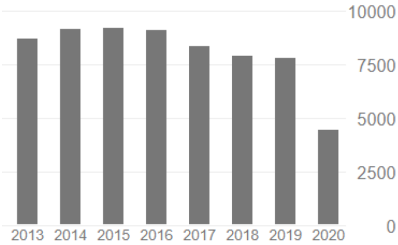
- Arguably, the most popular local image descriptor BCE, (vs SURF, ORB, Shape Context, HOG, etc).
- Published by D. Lowe, [Lowe, 1999; Lowe, 2004].
- Patented in Canada by the University of British Columbia.
- Defined upon the scale-space theory [Lindeberg, 1993].
- Exploits local partial derivatives of intensities (gray-scale).

| TITLE | CITED BY | YEAR |
|---|---|---|
| Distinctive image features from scale-invariant keypoints<br>DG Lowe<br>International journal of computer vision 60 (2), 91-110 | 58356 | 2004 |
| Object recognition from local scale-invariant features<br>DG Lowe<br>International Conference on Computer Vision, 1999, 1150-1157 | 19668 | 1999 |

# D. Lowe

Cited by VIEW ALL

| | All | Since 2015 |
|---|---|---|
| Citations | 108769 | 46866 |
| h-index | 52 | 35 |
| i10-index | 87 | 55 |



# SIFT pipeline

1. Build a scale-space image representation.
2. Detect PoI in scale and space (blobs).
3. Assign an orientation to each PoI.
4. Compute local descriptor (histogram of local orientations).
5. [optional] Image representation.

We already know:

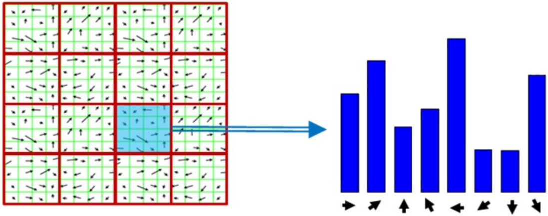
- Scale-space.
- Blob detection.
- Local orientation.

Let us focus on the local descriptor.

# Local descriptor

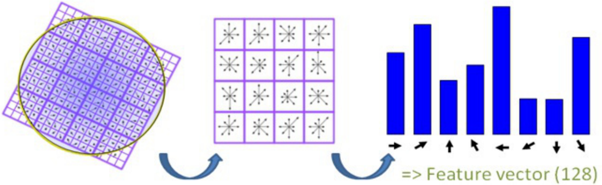
For each PoI ($p_i = [x_i, y_i, \theta_i, \sigma_i]$).

1. Take a $16 \times 16$ window around it (opt. window within $\sigma_i$).
2. Divide this region into 16 sub-blocks, e.g., of $4 \times 4$ pixels.
3. For each sub-block, create a 8 bins histogram of orientations. (local orientations are normalized w.r.t. $\theta_i$).
4. The final descriptor is a unit-normalized vector of length 128.



# Local orientation

As mentioned in step 3 from previous slide, local orientations are normalized w.r.t., a canonical orientation, i.e., the orientation of the PoI itself.




Think about it, the PoI's local orientation becomes $0$, and every other local orientation gets aligned accordingly.

- This makes the descriptor robust to rotation variations.

# Outline

# Intro

Binary images have little gradient information.

- Local orientation is not defined for every pixel.

Let us use only info from the shape's contour.

- Find contour.
- Sample point uniformly (PoI).
- Compute local orientation for each PoI.
- Define characteristic scale.

# Shape Context I

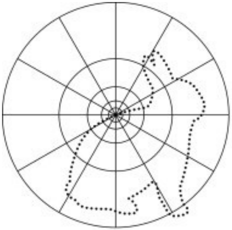
S. Belongie & J. Malik (2000). “Matching with Shape Contexts”.

## Local orientation:

Slope from PoI to nearest point.

## Characteristic scale:

Average pairwise distance between sampled points.



# Shape Context II

Use a log-polar grid (60 cells):

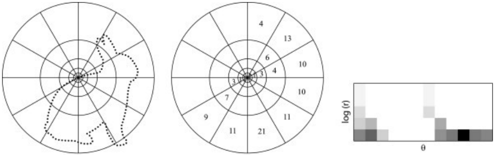
- 5 distance intervals $(0.125, 0.25, 0.5, 1, 2)\times$ charact. scale.
- 12 orientation intervals (every 30 degrees).

## Description:

For each PoI, count how many neighboring point are inside each cell of the log-polar grid.

## Rotation invariance:

Also align all point w.r.t. a canonical orientation.

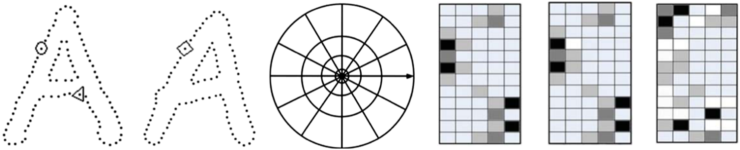



(a) (b) (c)

# Shape Context III

Each point is described by a vector of 60 elements (normalized).

There is one local descriptor per PoI.



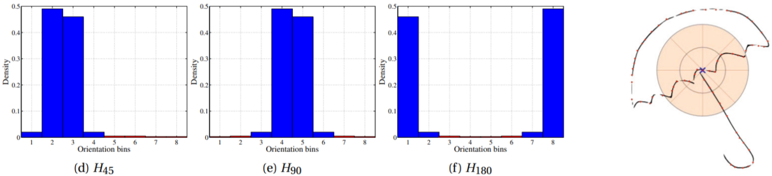
They can be treated as any other local descriptor, and used for:

- Matching.
- Alignment.
- Classification.
- etc.

# Histogram-of-Orientations Shape-Context, HOOSC

Roman-Rangel et al. 2010. 2014.

- Include orientation also for nearby points.
- Use linear-polar grid.
- Each polar cell becomes a histogram of orientations.
- Only 2 distance intervals at $(0.5, 1)\times$ char. scale.
- Descriptor is now of 128 elements.
- Efficient and good performance for retrieval.




(d) $H_{45}$ (e) $H_{90}$ (f) $H_{180}$

# Q&A

Thank you!

francisco.roman@itam.mx