Seminars

Universidad Politécnica de Madrid

Feature Extraction in Images by Keypoints

Prof. Raúl Alonso Calvo

ralonso@fi.upm.es

D-2315

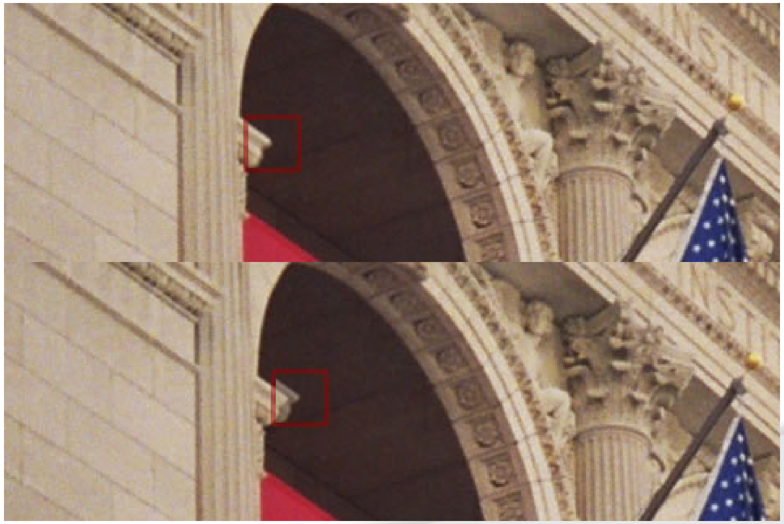
Tel: +34 91 336 37 48

Describe image by Keypoints

- One way of doing so is by applying the following steps:
  1. Compute distinctive keypoints in both images (for example, corners)
  2. Compare the keypoints between the two images to find matches
  3. Use the matches to find a general mapping between the images (for example, a homography)
  4. Apply the mapping on the first image to align it to the second image

Corner detectors: Harris corner detector

- The Harris Corner Detector is a mathematical operator that finds features in an image:
  - find little patches of image (or "windows") that generate a large variation when moved around
  - A small movement in a set of pixels has a great variance in the set



http://aishack.in/tutorials/harris-corner-detector/

Corner detectors: Harris corner detector

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]^2$$

- E is the difference between the original and the moved window.
- u is the window's displacement in the x direction
- v is the window's displacement in the y direction
- w(x, y) is the window at position (x, y). This acts like a mask. Ensuring that only the desired window is used.
- I is the intensity of the image at a position (x, y)
- I(x+u, y+v) is the intensity of the moved window
- I(x, y) is the intensity of the original

- Taylor series and simplification we get:

$$E(u,v) \approx [u \quad v]\, M \begin{bmatrix} u \\ v \end{bmatrix}$$

$$M = \sum w(x,y) \begin{bmatrix} {I_x}^2 & I_x I_y \\ I_x I_y & {I_y}^2 \end{bmatrix}$$

Corner detectors: Harris corner detector

- A score, R, is calculated for each window:

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

$$\det M = \lambda_1 \lambda_2$$

$$trace\, M = \lambda_1 + \lambda_2$$

- All windows that have a score R greater than a threshold are corners
  - good tracking points

Corner detectors: Shi-Tomasi detector

- Based on Harris corner detector, changes the R function

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

$$\det M = \lambda_1 \lambda_2$$

$$trace\, M = \lambda_1 + \lambda_2$$

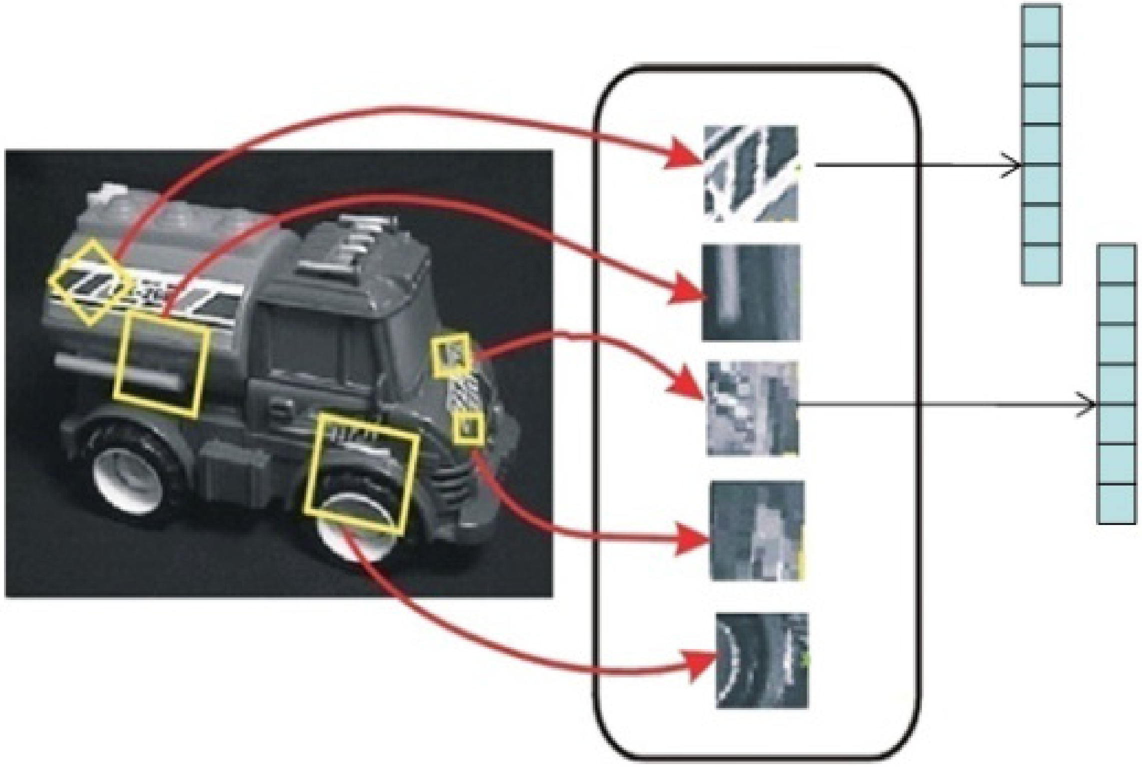
$$\Rightarrow \quad R = min(\lambda_1, \lambda_2)$$

- It is faster and performs very well for detecting keypoints

Patch descriptors

- A descriptor is some function that is applied on the patch to describe it in a way that is invariant to all the image changes that are suitable to our application (e.g. rotation, illumination, noise etc.).
- A descriptor is “built-in” with a distance function to determine the similarity, or distance, of two computed descriptors. So to compare two image patches, we’ll compute their descriptors and measure their similarity by measuring the descriptor similarity, which in turn is done by computing their descriptor distance.

Patch descriptors



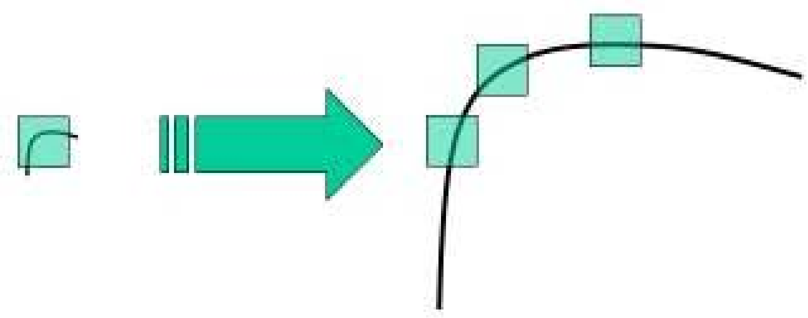
- SIFT & SURF &

SIFT: Scale-invariant feature transform

- Algorithm to detect and describe local features in images
- The algorithm was published by David Lowe in 1999
- The algorithm is patented in the US
  - the owner is the University of British Columbia
- The SIFT algorithm extract keypoints and compute its descriptors
- It is rotation-invariant
  - if the image is rotated, we can find the same corners
  - corners remain corners in rotated image also

Detection of scale-space extrema

- We can't use the same window to detect keypoints with different scale. Larger corners require larger windows.



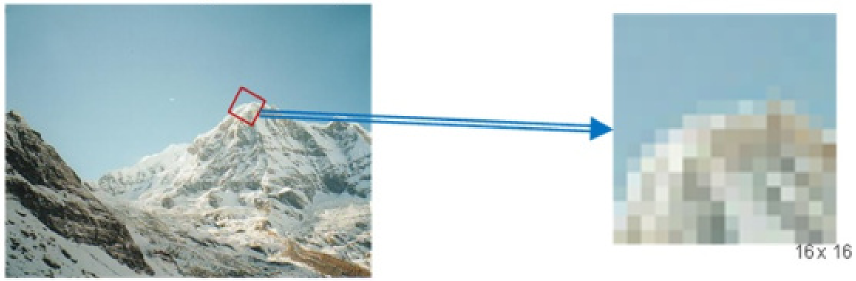
- SIFT algorithm uses Difference of Gaussians which is an approximation of Laplacian of Gaussian (LoG)
- We can find the local maxima across the scale and space

Local extrema detection - Keypoint Localization

- Once potential keypoints locations are found, they have to be refined to get more accurate results
- if the intensity at this extrema is less than a threshold value (0.03 as per the paper), it is rejected
- DoG has higher response for edges, so edges also need to be removed
  - used a 2x2 Hessian matrix (H) to compute the pricipal curvature
  - If this ratio is greater than a edgeThreshold in keypoint is discarded

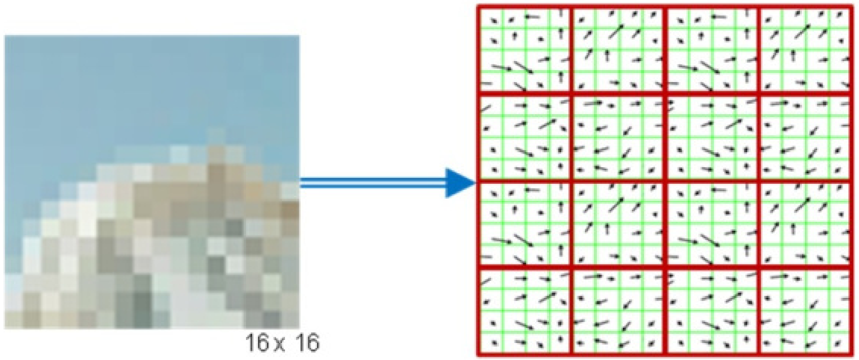
Orientation Assignment

- A neigbourhood is taken around the keypoint location depending on the scale, and the gradient magnitude and direction is calculated in that region.
- An orientation histogram with 36 bins covering 360 degrees is created

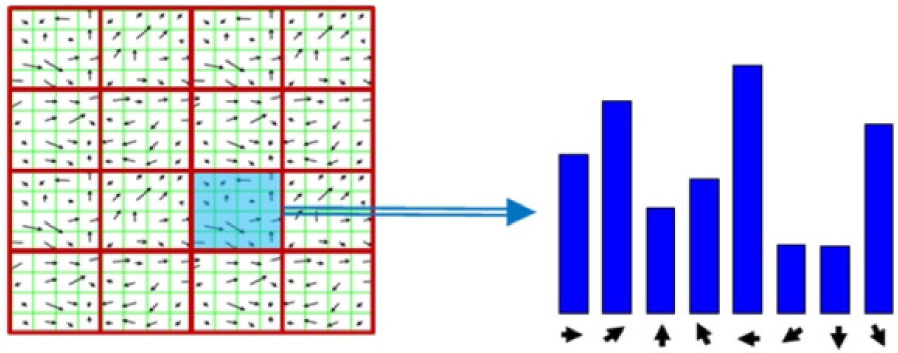


Keypoint descriptor

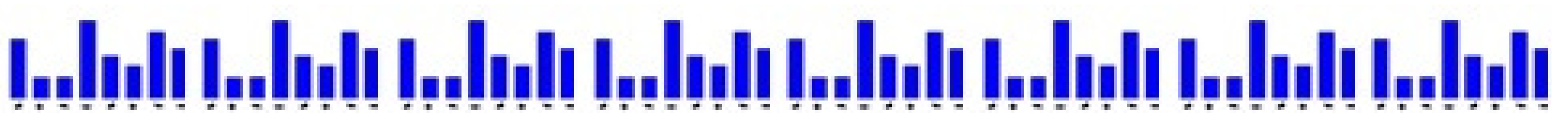
- A 16x16 neighbourhood around the keypoint is taken
- It is devided into 16 sub-blocks of 4x4 size
- For each sub-block, 8 bin orientation histogram is created
- A total of 128 bin values are available.



Keypoint descriptor



- Thus, descriptor for a keypoint is like:



- 28 (16*8) dimensional feature vector

Keypoint matching

- Keypoints between two images are matched by identifying their nearest neighbours
- In some cases, the second closest-match may be very near to the first.

BRIEF: Binary Robust Independent Elementary Features

- Binary descriptors are composed of three parts
  - A sampling pattern
  - orientation compensation
  - sampling pairs
- Consider a small patch centered on a keypoint. We'd like to describe it as a binary string
- Now go over all the pairs and compare the intensity value of the first point in the pair with the intensity value of the second point in the pair
  - If the first value is larger then the second, write '1' in the string, otherwise write '0'

Other binary descritors

- Puzzle task:
  - ORB
  - BRISK
  - Freak

References

- Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of Alvey Vision Conference, pp. 147-151 (1988).
- Lowe, David G. (1999). "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. pp. 1150–1157. doi:10.1109/ICCV.1999.790410.
- Lowe, David G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints". International Journal of Computer Vision 60 (2): 91–110. doi:10.1023/B:VISI.0000029664.99615.94.
- Herbert Bay , Andreas Ess , Tinne Tuytelaars , Luc Van Gool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding, v.110 n.3, p.346-359, June, 2008 [doi>10.1016/j.cviu.2007.09.014]

References

- Edward Rosten , Tom Drummond, Machine learning for high-speed corner detection, Proceedings of the 9th European conference on Computer Vision, May 07-13, 2006, Graz, Austria [doi>10.1007/11744023_34]
- Calonder, Michael, et al. "Brief: Binary robust independent elementary features." Computer Vision–ECCV 2010. Springer Berlin Heidelberg, 2010. 778-792.
- Rublee, Ethan, et al. "ORB: an efficient alternative to SIFT or SURF." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.
- Leutenegger, Stefan, Margarita Chli, and Roland Y. Siegwart. "BRISK: Binary robust invariant scalable keypoints." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.

References

- Alahi, Alexandre, Raphael Ortiz, and Pierre Vandergheynst. "Freak: Fast retina keypoint." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.
- http://docs.opencv.org/
  - http://docs.opencv.org/3.1.0/db/d27/tutorial_py_table_of_contents_feature2d.html
- https://gilscvblog.com
  - https://gilscvblog.com/2013/08/18/a-short-introduction-to-descriptors/