Master in Computer Vision Barcelona

Project Module 1

Content based image retrieval

Coordination: Ramon Morros / Verónica Vilaplana

ramon.morros@upc.edu, veronica.vilaplana@upc.edu

Week 4

W3 Homework

Query results evaluation

QST1 (Simple)

Team	method1	map@1
Team5	method1	**0.960**
Team1	method1	0.940
Team7	method1	0.940
Team9	method1	0.920
Team4	method1	0.860
Team2	method1	0.840
Team8	method1	0.840

QST2 (Complex)

Team	method1	map@1
Team8	method1	**0.795**
Team4	method1	0.769
Team5	method1	0.744
Team7	method1	0.641
Team9	method1	0.538
Team2	method1	0.436
Team1	method1	0.385

W3 Homework

QST1 (Simple)

Query results partial evaluation

None

Team	method1	map@1
Team9	method1	1.000
Team1	method1	0.950
Team5	method1	0.950
Team7	method1	0.900
Team8	method1	0.900
Team2	method1	0.850
Team4	method1	0.850

Noise

Team	method1	map@1
Team5	method1	1.000
Team7	method1	1.000
Team1	method1	0.833
Team4	method1	0.833
Team8	method1	0.750
Team9	method1	0.750
Team2	method1	0.667

Color change

Team	method1	map@1
Team1	method1	1.000
Team2	method1	0.944
Team5	method1	0.944
Team7	method1	0.944
Team9	method1	0.944
Team4	method1	0.889
Team8	method1	0.833

W3 Homework

Text evaluation

QST1 (Simple)

Team	method1	Avg. Text Distance	Total lines	Checked lines
Team2	method1	**1.25**	50	40
Team8	method1	1.43	50	44
Team4	method1	1.88	50	50
Team7	method1	2.05	50	37
Team5	method1	2.62	50	50
Team1	method1	3.52	50	42
Team9	method1	6.72	50	50

Team	Method	Avg. Text Distance	Total lines	Checked lines
Team8	method1	**1,39**	39	36
Team4	method1	1,58	39	38
Team1	method1	2.83	39	29
Team7	method1	3.62	39	26
Team2	method1	4,21	39	38
Team5	method1	4.64	39	39
Team9	method1	7,18	39	38

```
from Levenshtein import distance
for ii in range(len(text_lines)):
    current_dist = distance(text_lines[ii], hypo_text_lines[ii])
    total_dist = total_dist + current_dist
avg_dist = total_dist/len(text_lines)
```

Average Levenhstein distance. Only the valid lines (the files wich contain text) are considered.

W3 Homework

Text box evaluation

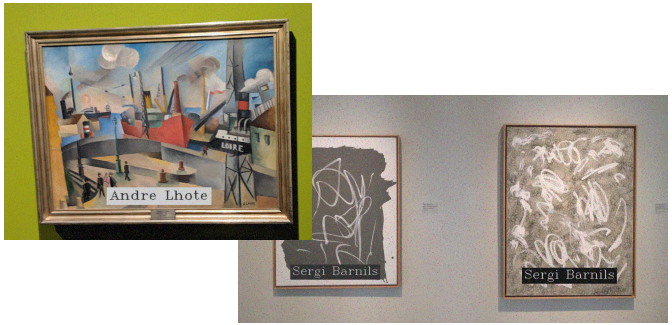
QST1 (Simple)

Team	method1	Precision	Recall	F1	IoU
Team2	method1	0.72	0.72	0.72	**0.93**
Team9	method1	0.62	0.62	0.62	**0.93**
Team7	method1	0.68	0.68	0.68	0.86
Team4	method1	0.86	0.86	0.86	0.79
Team5	method1	0.86	0.86	0.86	0.79
Team8	method1	**0.88**	**0.88**	**0.88**	0.79
Team1	method1	0.78	0.78	0.78	0.73

QST2

Team	**method1**	**Precision**	**Recall**	**F1**	**IoU**
Team7	method1	0.08	0.08	0.08	**0.95**
Team9	method2	0.58	0.56	0.57	0.92
Team2	method1	0.56	0.62	0.59	0.88
Team4	method1	0.90	0.90	0.90	0.81
Team8	method1	0.95	0.92	**0.94**	0.81
Team1	method1	0.74	0.67	0.70	0.79
Team5	method1	0,80	0.82	0.81	0.78

W4 Datasets



Museum datasets

- [Can Framis Museum]
- [Figueres 120 years expo]
- [Kode Bergen]

Query dataset (development with GT and test without GT)

- Original and paintings with
  - superimposed text on a **semitransparent** box (painter name, different fonts, sizes and positions)
  - noise (some random samples)
  - color changes (some random samples, random Hue changes)
  - double queries
  - **The query set can contain images not in the database**

- QSD1-W4 (30) / QST1-W4 (50) pictures with background, with overlapping text (name of painter), **one, two or three** paintings per image, some paintings with noise, some paintings with changes in color, some paintings not in database

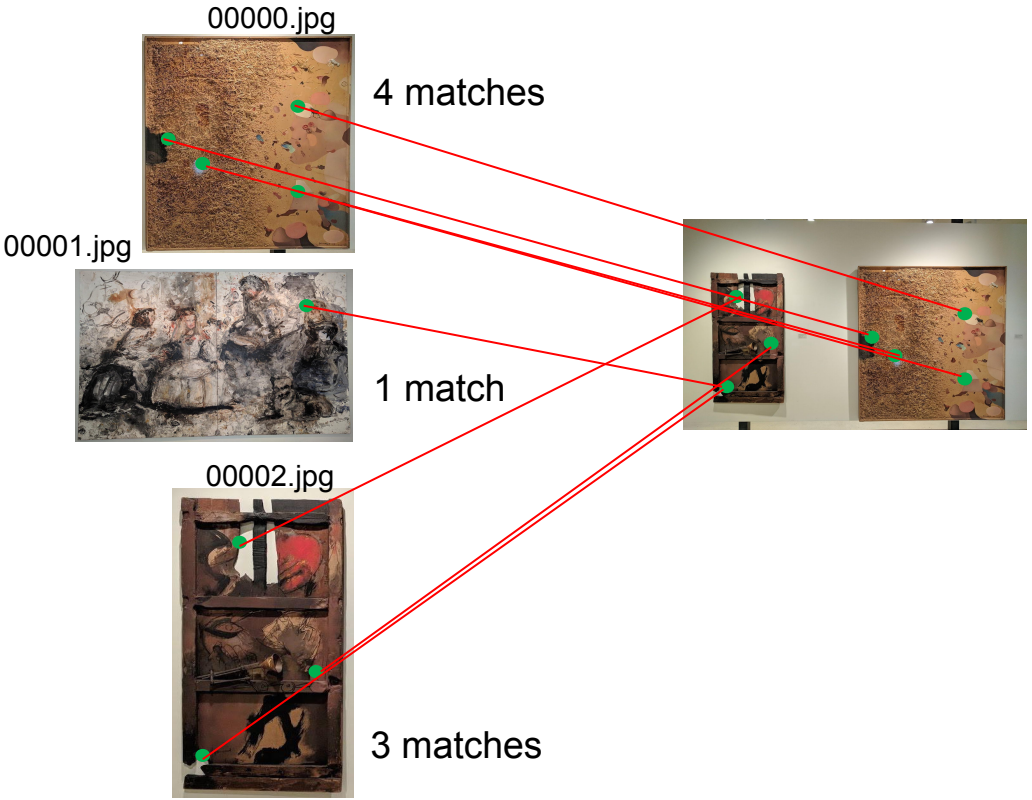
W4 – Content based image retrieval

- **Goal:** search images from a large image database (DB) based on visual contents

- **Similarity:** based on number of matches between local descriptors
  - Match features. Discard false/ambiguous matches
  - Select DB images with larger number of matches
  - Discard images with ‘small’ number of matches

- **CBIR steps:**
  - Index the DB: generate keypoints, descriptors (e.g. SIFT) for all images
  - Extract keypoints, features from query image (e.g. SIFT)
  - Compute matches between descriptors of keypoints from the query image and each DB image descriptor, and order the DB images according to the number of matches

W4 – Query method



[[**2**,7,9,...],[**0**,23,44,...]]

W4 – Query method

00000.jpg

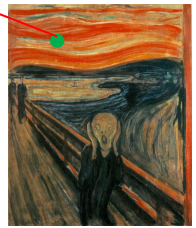
1 matches

00001.jpg

0 match

00002.jpg

0 matches



[[-1]]

W4 - Tasks

- Given the museum and the query dataset. For each image in the query dataset, retrieve the K most similar images in the Museum dataset, ordered by score

- Image descriptor: keypoints + local descriptors

Task 1: Detect keypoints and compute descriptors in Museum and query images

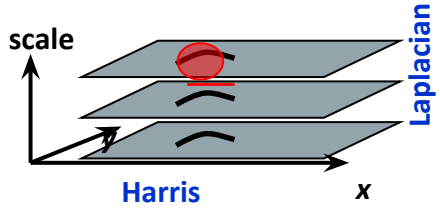
Task 2: Find tentative matches based on similarity of local appearance and verify matches

Task 3: Evaluate the system on QSD1-W4, map@k

Task 4: Evaluate best system from previous week on QSD1-W4

W4 - Task1

- **Keypoint detection**

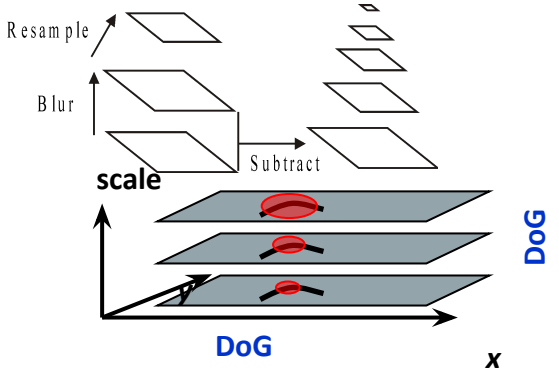


Rotation, translation, intensity invariant

- **Harris corner detector:** uses the autocorrelation (second moment) matrix .
  Important difference in all directions-> interest point [Harris ‘88]

Scale invariant

- **Harris Laplacian** [Mikolajczyk & Schmid ‘01]
  - Find local maxima/minima of:
    - Harris corner detector in space (image coordinates)
    - Laplacian in scale
- **Difference of Gaussians (DoG or SIFT)** [Lowe’99]
  - Find local maxima/minima of
    - Difference of Gaussians in space and scale

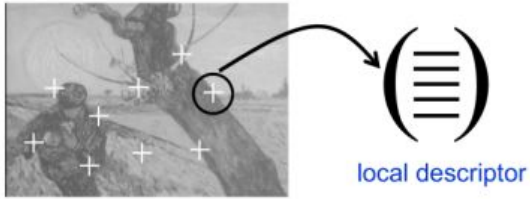


W4 - Task 1

- **Local descriptors:**

  Tons of options!

  - SIFT
  - SURF
  - ORB (FAST+BRIEF)
  - LBP
  - HOG
  - PCA-SIFT
  - Color-SIFT
  - GLOH
  - DAISY
  - …



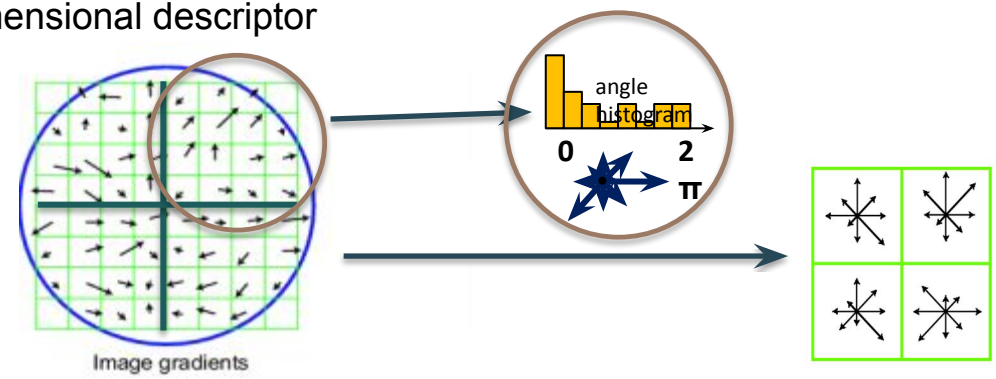
https://docs.opencv.org/master/db/d27/tutorial_py_table_of_contents_feature2d.html

W4 - Task 1

- SIFT descriptor
  - Take 16x16 window around detected keypoint (8x8 shown below), rotated along keypoint direction
  - Compute edge orientation (angle of the gradient – 90°) for each pixel
  - Throw out weak edges (threshold gradient magnitude)
  - Divide the 16x16 window into a 4x4 grid of cells (2x2 case shown below)
  - Create histogram (8 bins) of surviving edge orientations for each cell
  - 16 cells * 8 orientations = 128 dimensional descriptor
  - Normalize vector

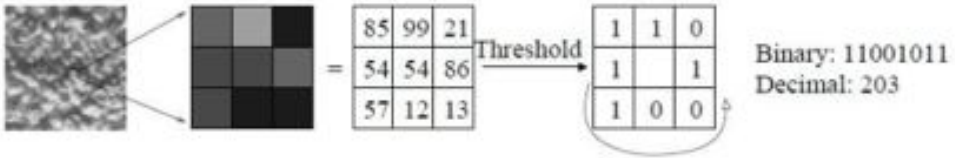
```
pip install opencv-contrib-python==4.4.0.46

import cv2
kp, desc = sift.detectAndCompute(gray,None)
```

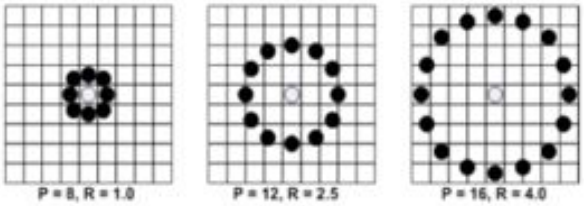


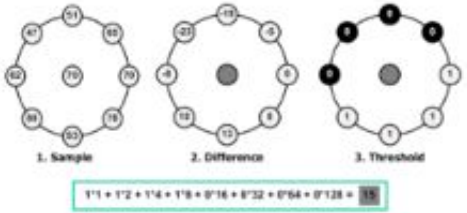
W4 - Task 1

- LBP (Local Binary Pattern) descriptor
  - Divide image into blocks, for each pixel in the block compare to its 8 neighbors



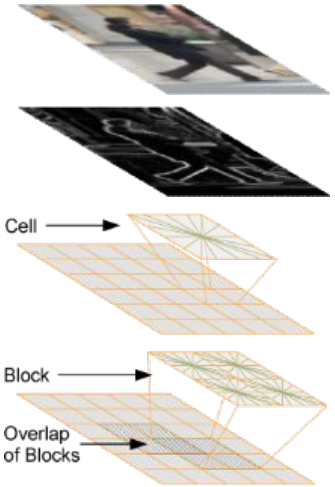
  - If the center pixel's value is greater than the neighbor's value, write 0, otherwise, write 1. The result is a 8-digit binary number. Compute histogram, over the block, of the frequency of each number
  - **Multiscale LBP:** different local neighborhoods. Bilineal interpolation is performed for points which are not centered on pixel

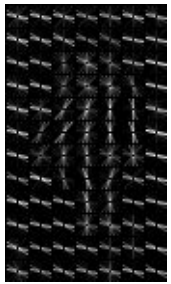


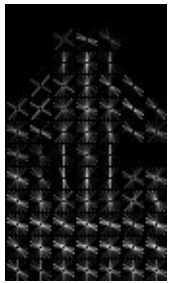


W4 - Task 1

- HOG Histogram of oriented gradient
  - Compute gradients on an image
  - Compute histograms of gradient orientation on ‘cells’ of typically 8x8 pixels
  - Normalize histograms within overlapping blocks of cells (typically 2x2 cells)
  - Concatenate histograms







W4 - Task 2

- Find tentative matches based on similarity of local appearance and verify matches
  - consider different similarity metrics

- Implement a system to discard queries not in the data set (unknowns)
  - Usually, a threshold in the number of matches
  - Define the threshold by optimizing F1 measure on the development set

W4 - Task 3

- Evaluate system based on keypoint descriptors on QSD1-W4

W4 - Task 4

- Optional: Evaluate your best query system **from previous week** on QSD1-W4

W4 - Submissions

- For each query, a list of the K best results (K=10).
  - **[[-1]] if ima not in dataset**
  - Only best method!
- For each query image, a text file with the text transcription (one line for each painting)

Note: Deliver files to:
/home/dlcv0X/m1-results/week4/QST1/**method1**/result.pkl
/home/dlcv0X/m1-results/week4/QST1/**method1**/text_boxes.pkl
/home/dlcv0X/m1-results/week4/QST1/**method1**/*.txt

- Tests sets delivered on Sunday 06 Nov 2021 at 14h
- Submit progress slides
  - Deadline slides: Sunday 06 Nov 2022 at 19:00
  - Deadline results: Sunday 06 Nov 2022 at 19:00
  - Deadline questions to teams: Monday 07 Nov 2022 at 14:00