Museum painting retrieval

Team 8

José Manuel López Camuñas, Marcos V. Conde, Alex Martín Martínez

15 / 11 / 2021

Summary of previous weeks

Week 1

- Color based retrieval: RGB 1D Histogram with Cosine Similarity
- Background removal: Otsu’s method with a binarization

Week 2

- Color based retrieval: Multi-block 3D RGB + Hellinger distance
- Text removal: Combination of Morphological Filters(tophat, blackhat,...)
- Background removal: Otsu’s + Morphological Operations (e.g. closing)

Summary of previous weeks

Week 3

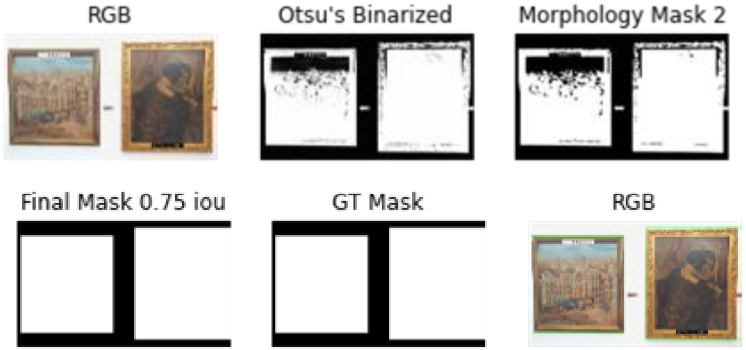
- Text based retrieval: Previous Text Detection + Levenshtein matching.
- Texture based retrieval: HoG and LBP
- Denoising: NL-means, Median Filter
- Combination: HoG + Multi-histogram RGB
- Background removal: Otsu’s + Morphological Operations (e.g. closing)

Week 4

- Text detection: Harris corner detector + Morphological operators
- Keypoint based retrieval: SIFT and ORB
- Denoising: NL-means, Median Filter
- Combination: HoG + SIFT
- Background removal: Otsu’s + Morphological Operations (e.g. closing)

Current Retrieval Pipeline

1. Input image → 2. Denoising → 3. Background removal and rotation correction

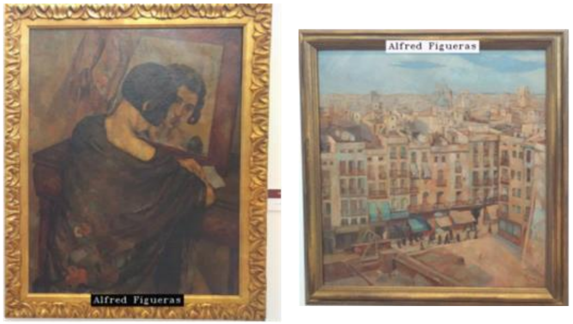


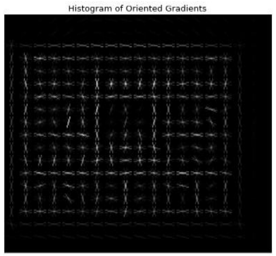
Current Retrieval Pipeline

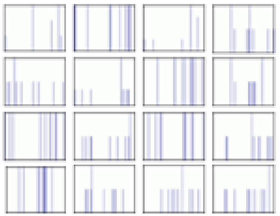
4. Obtain single paintings from the photo.

5. Feature Extraction + Similarity Search

6. Image Retrieval







Task 1: Background removal and rotation

We use the same background removal method as previous weeks for a fair comparison.

Task 1: Background removal and rotation

The rotation angle is estimated as follows:

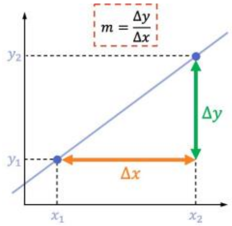
1) The RGB image is converted to GrayScale.

1) We use the gradient-based **Canny Filter** for edges detection. The *cv2.Canny* implementation using default parameters and *apertureSize=3.*

1) We use the **Hough Line Transform** to detect straight lines form the edges previously obtained [1]. We filter lines such that the minimum length is 50.

1) We **calculate the angle of each line** using the equation below, and converting the result from radians to degrees. Finally, we use the mean of all the obtained angles < 30º, as the final rotation angle.

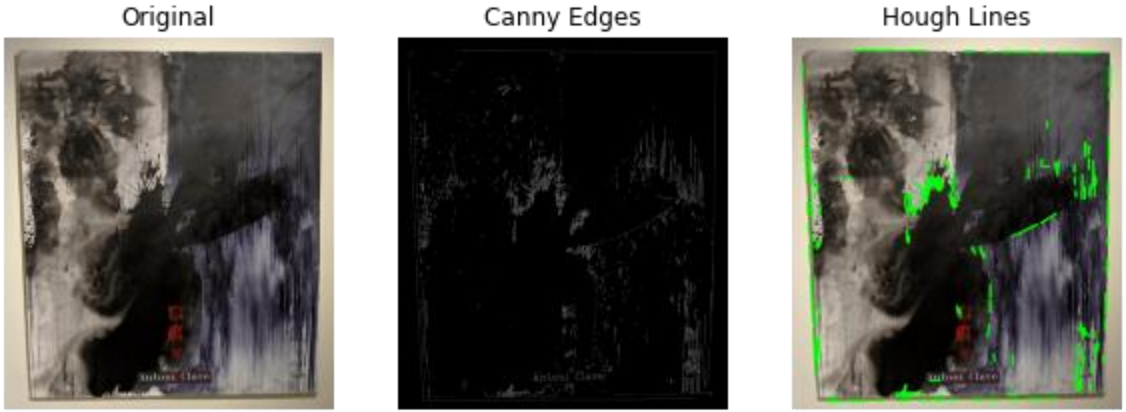
$$\arctan\left((y_2 - y_1)/(x_2 - x_1)\right)$$

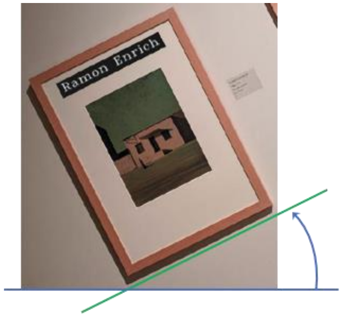


[1] https://docs.opencv.org/3.4/d9/db0/tutorial_hough_lines.html

Task 1: Background removal and rotation

The filter of considering only angles < 30º, helps to “remove” the vertical lines and non-realistic orientations.
We consider the Mean Angular Error not very robust to outliers. Despite our MAE is 0.897 we consider the Median Angular Error a better metric, achieving 0.497.

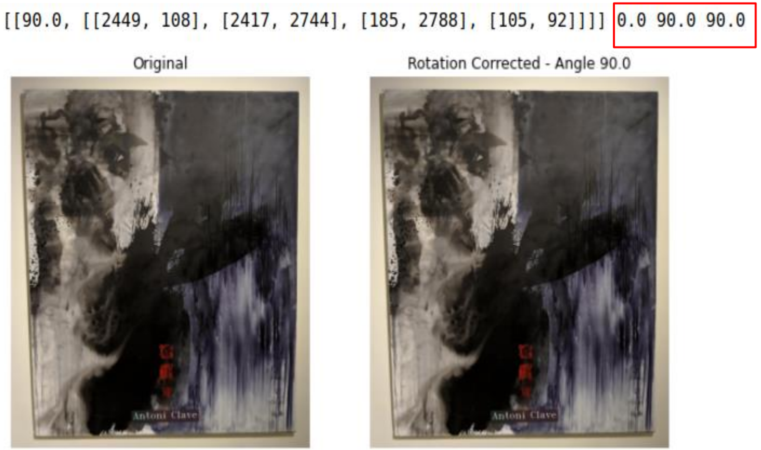




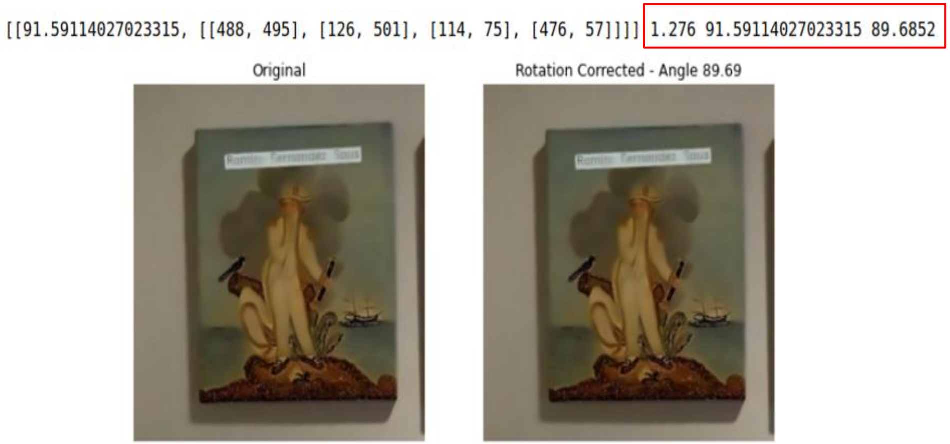
alpha = 26º

Task 1: Background removal and rotation

The method could determine the rotation angle of the image with precision. The numbers in the red boxes are: the error, the actual angle, and the estimated angle.



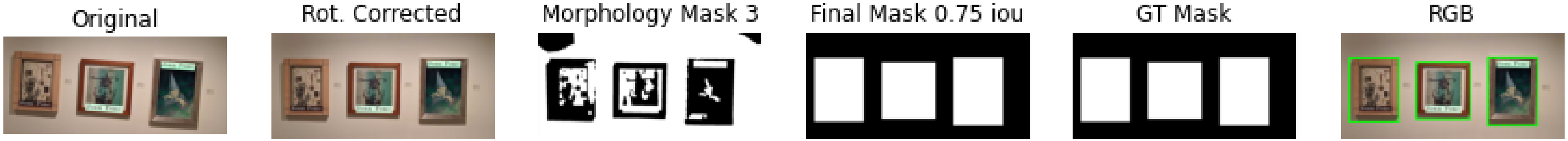
Original Angle = 90º
Predicted Angle = 90º
Error = 0



Original Angle = 91.6º
Predicted Angle = 89º
Error = 1.276

Task 1: Background removal and rotation

Pipelines for background removal and rotation correction



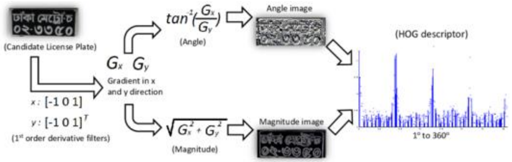
Final crops extracted from the image

Task 2: Retrieval

- Tested methods for retrieval:
  - HoG + Multi-histogram
  - SIFT
  - ORB
  - SIFT + ORB
  - SIFT + HoG + Mult-histogram

Task 2: Retrieval

- HoG + Multi-histogram(linear combination of similarities)



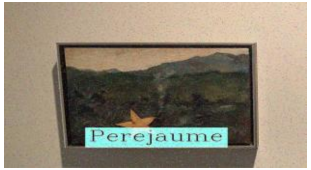
Compute similarities with the Database

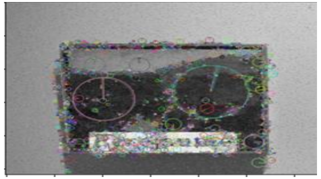
$$S = \alpha * T_{score} + \beta * C_{score}$$

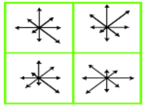
Where **α** and **β** are selected empirically to optimize the metric MAP and both T and C scores are the similarities obtained using Texture (HOG) and Color (Multi-Histogram)

Task 2: Retrieval

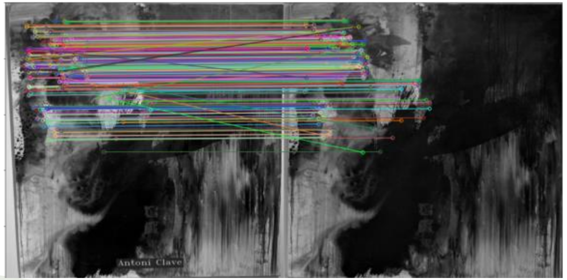
- For SIFT and ORB the usage of more keypoints (up to 1000), higher resolution images (from 300x300 to 500x500) and thresholding for missing query detection improved the method.

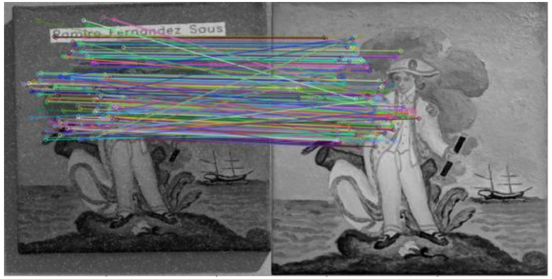






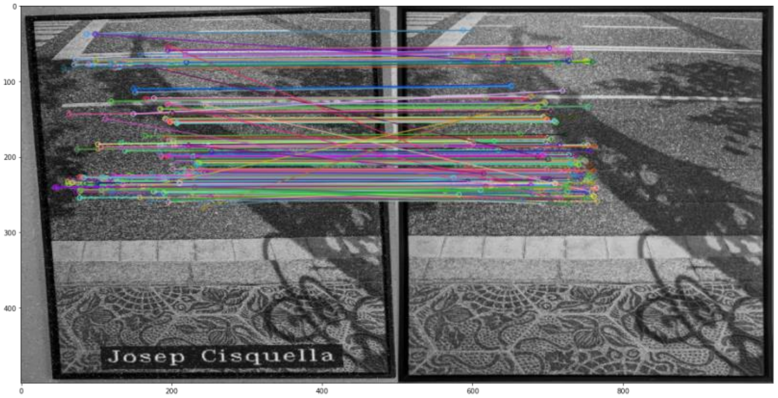
Descriptors to match images

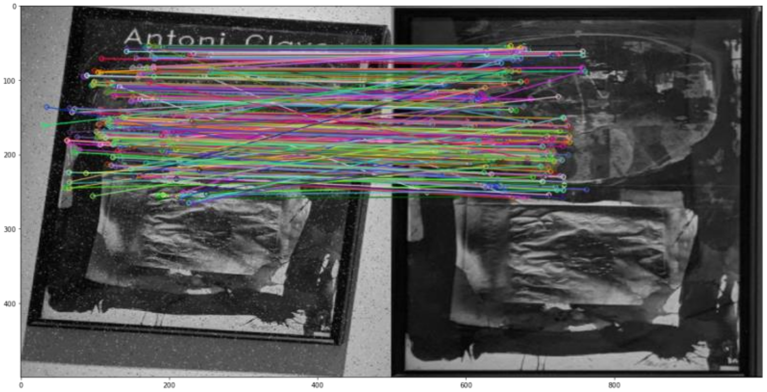




Task 2: Retrieval

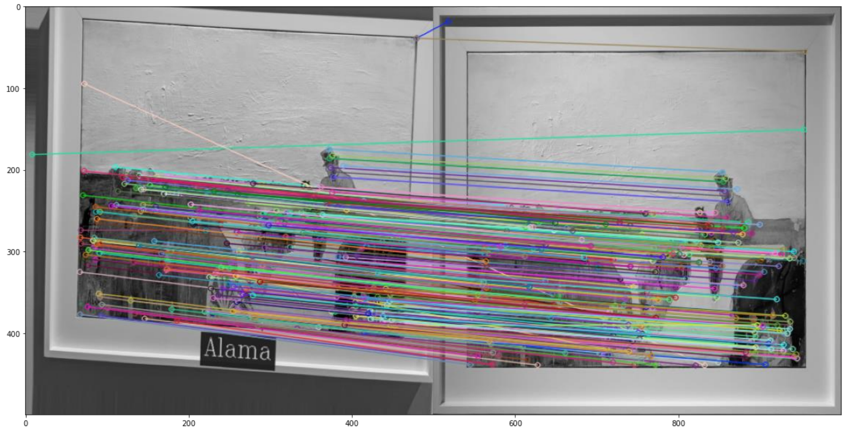
- QSD1-W5 Examples

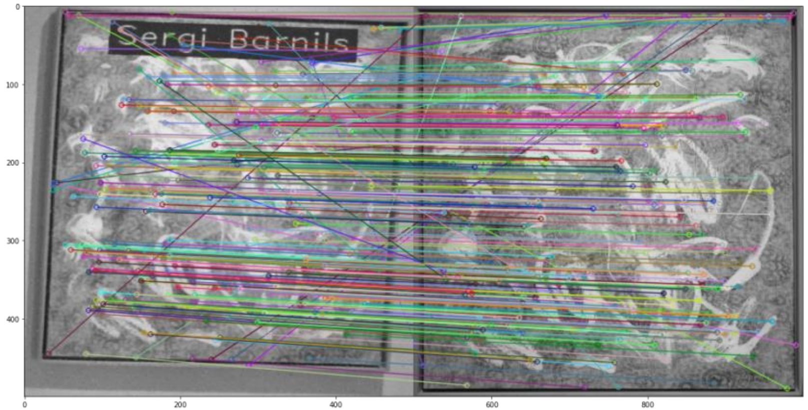




Task 2: Retrieval

- QST1-W5 Examples



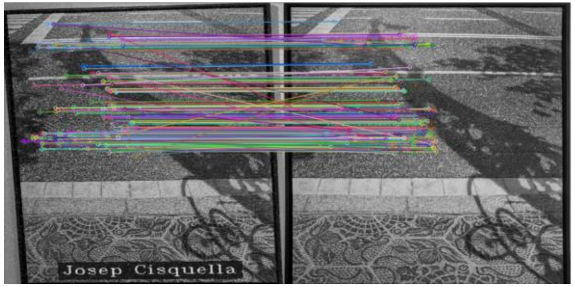


Task 2: Retrieval

When combining the keypoint descriptors we used the following steps:

1) Detecting keypoints
2) Normalization with respect to the greatest number of matches detected and use this as a similarity value.
3) Aggregation of the similarities.

$$S = \alpha * T_{score} + \beta * C_{score}$$



Matrix with number of keypoints matched

Similarity value matrix

Task 2: Retrieval

Some images from the Query set may not be included in the provided database, we need to recognize them:

- How? Set a certain threshold that determines when a query is missing.
- 1st idea : Manual thresholding(by observing similarity matrix).
- 2nd idea: Using Mean and Standard Deviation (mean+-std).
- Final idea: Perform Grid-Search from 0.01 to 1.00 with 0.01 step.

We obtain the optimal threshold for each type of feature representation!

Task 3: Evaluation of retrieval

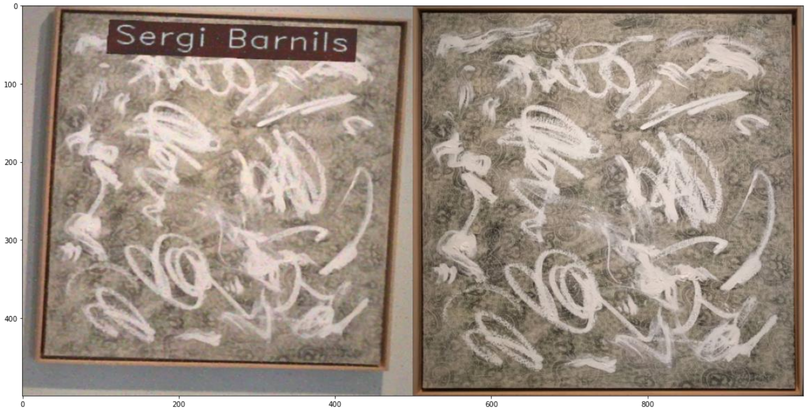
	MAP@10	MAP@5
HoG+Multi-histogram(α=0.7, β=0.3, th = 0.62)	0.5333 (16/30)	0.4944 (14/30)
SIFT (th = 0.29)	**0.7333 (22/30)**	**0.7130 (21/30)**
ORB (th = 1.00)	0.3333 (10/30)	0.3333 (10/30)
SIFT+ORB(α=0.7, β=0.3, th=0.19)	0.7333 (22/30)	0.7130 (21/30)
SIFT+ (HoG+Multi-histogram) (α=0.7, β=0.3, th=0.38)	0.7333 (22/30)	0.7296 (21/30)

Task 3: Evaluation of retrieval

- Qualitative Results: QSD1-W5

Task 3: Evaluation of retrieval

- Qualitative Results: QST1-W5



Task 4: Museum exhibition

- Main objective: Group paintings by visual perception features.

- Features that influence in a paint:
  - Edges
  - Palette of colors
  - Saturation
  - Light
  - Contrast
  - Shapes
  - Dimension/size/resolution

Task 4: Museum exhibition

- We use the following features to obtain k=10 clusters from the data:

  - Edges → Laplacian
  - Palette of colors → Hue
  - Saturation → Saturation
  - Light → Value

We use the mean and standard deviation of such features.

We also considered to use the Chroma-Luma features from the YUV representation instead of HSV.

Task 4: Museum exhibition

- Calculate mean and standard deviation for the Hue, Saturation, Value, and the Laplacian of the image.

  ➡ Descriptor 8-dimensional for each image. (287,8)

- Principal Component Analysis: To visualize in a 2D plane the clusters (or rooms) , we use PCA to project into 2 components the 8-dimension descriptor.

  We use `sklearn.decompositon.PCA()`.

  ➡ 8D → 2D (287,2)

Task 4: Museum exhibition

With the array obtained after the PCA we run the Gaussian Mixture Model to generate the k(10) clusters.

The gaussian mixture model will try to maximize the maximum likelihood through the parameters of the k gaussian distributions, μ,N, Σ, and π which will depend on γ.

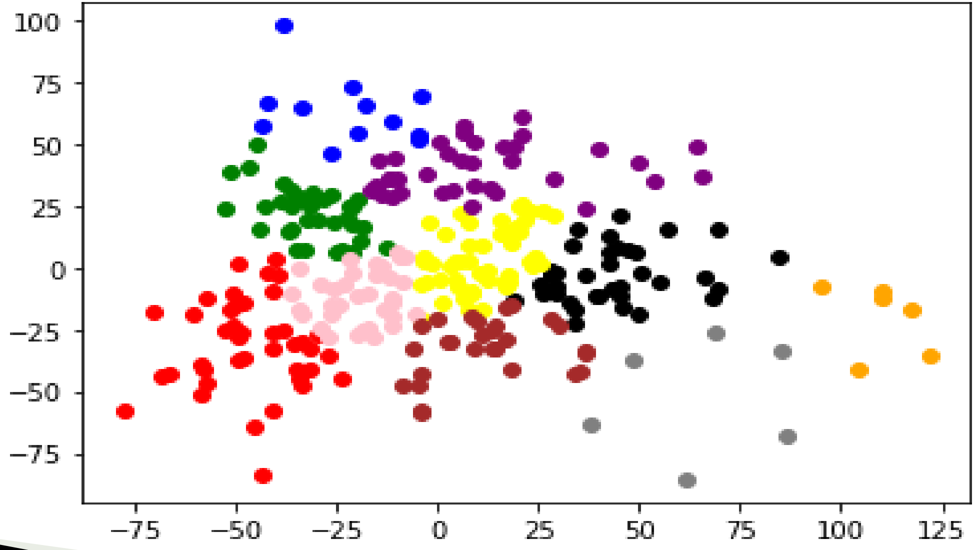
$$\gamma_k(x_i) = \frac{\pi_k N(x_i \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j N(x_i \mid \mu_j, \Sigma_j)}$$

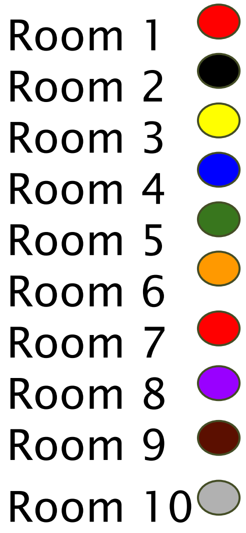
Recursive procedure:

1. Start: Given K, provide the initial parameters: $\{\pi_k, \mu_k, \Sigma_k\}, \quad k = 1 \ldots K$
2. Classify: given $\{\pi_k, \mu_k, \Sigma_k\}$ compute the $\{\gamma_k\}$
3. Re-center: Given $\{\gamma_k\}$, compute the $\{\pi_k, \mu_k, \Sigma_k\}$
4. Repeat 2-3 until convergence

Task 4: Museum exhibition

After the iterative process each point from the feature space will be labeled into one of the 10 groups.





Task 4: Museum exhibition

Undetailed

Task 4: Museum exhibition

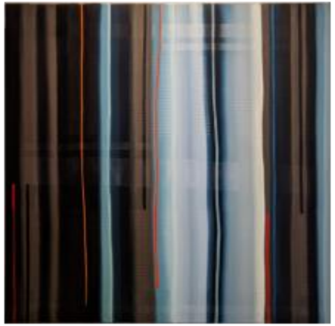
Light and ocre

Task 4: Museum exhibition

Gray levels

Task 4: Museum exhibition

Contrasts



Task 4: Museum exhibition

Lugubrious

Task 4: Museum exhibition

Happy blue

Task 4: Museum exhibition

Randomness

Task 4: Museum exhibition

Colorful

Task 4: Museum exhibition

Many things

Task 4: Museum exhibition

Darkness





Discussion and Results

- The complete retrieval pipeline allows us to retrieve the closest images from our DataBase given an input image captured “in the wild”.

- The pipeline is robust against illumination changes, noise, rotations and camera pose, this is because we use SIFT invariant keypoints as our main descriptor.

- Pure interpretable image processing features sometimes are more suitable than learnable features like CNNs.

- Other techniques such as Query Expansion or Database-side Augmentation (DBA) can improve the retrieval performance without modifying the descriptors.