A retina-inspired descriptor for image classification

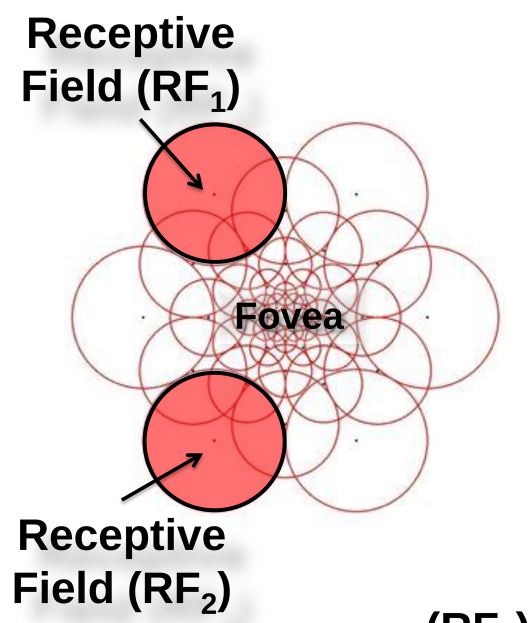
Cristina Hilario[1], Kartheek Medathati[2], Pierre Kornprobst[2], Vittorio Murino[1] and Diego Sona[1]
[1] Pattern Analysis and Computer Vision, PAVIS, Istituto Italiano di Tecnologia, Genova, Italy
[2] Neuromathcomp, INRIA, Sophia Antipolis, France

Abstract

- Retina **extracts many features** from visual stimuli with a high efficiency. Our **goal** is to build a **bio-inspired descriptor for scene categorization based on low-level processing** performed by the **retina.**
- We propose a **new image descriptor** based on **the center-surround organization of the ganglion receptive fields**. Each receptive field is described with a **linear-nonlinear model (LN)** taking into account the inhibitory surrounds.
- **FREAK[1] descriptor** is used as baseline, which is enriched sticking to biological data and models of retina. We define a **set of FREAK-variant descriptors** and test them on the problem of scene classification.

Proposed Method

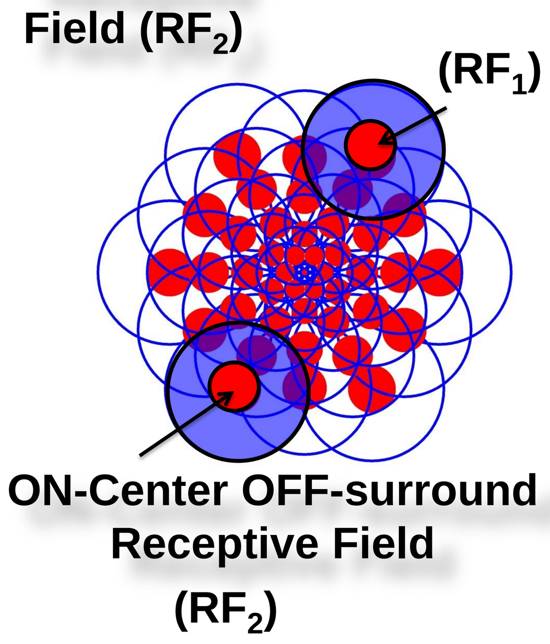
1. Retinal Sampling Pattern



FREAK:

- **Circular grid** of concentric distribution of overlapping receptive fields (RFs)
- **Higher density** of RFs near the **fovea**
- Considers the difference in intensity between **pairs of receptive fields**:

$$binary\ descriptor = \begin{cases} 1 & \text{if } (intensity(RF_1) - intensity(RF_2)) > 0 \\ 0 & \text{otherwise} \end{cases}$$



Our model:

- **Circular center-surround** organization of RFs; the radius of the **surround** is calculated using the retinal eccentricity[2] ρ:

$$r_S(\rho) = 0.203\rho^{0.472}$$

- **ON-center OFF-surround** processing is mimicked using a **difference of Gaussians (DoG):**

$$K(x,y) = w_C G_{\sigma_C}(x,y) - w_S G_{\sigma_S}(x,y)$$
$$\sigma_S = 3\sigma_C$$

2. Approach

- **Retinal activity** is defined by a **linear-nonlinear (LN) model**, where the **activity A of a cell** is defined by:

$$A = f(RF) \quad \text{where} \quad RF\text{: Receptive field of the cell}$$
$$f\text{: Nonlinear function}$$

- Each **binary descriptor** is defined as the **difference in activity** between **pairs of RFs:**

$$binary\ descriptor = \begin{cases} 1 & \text{if } f(RF_1) - f(RF_2) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Descriptor 1: Consider the activity in the center of the RF

$$A = f(I * K_C(x,y)) \quad \text{where} \quad K_C(x,y) = w_C G_{\sigma_C}(x,y)$$
$$I\text{: Static Image}$$

$$binary\ descriptor = \boxed{0101100100\ldots1}$$

Center RF

Descriptor 2: Add the surround response (sign of the DoG)

$$A = sign(I * K(x,y)) \text{ where } K(x,y) = w_C G_{\sigma_C}(x,y) - w_S G_{\sigma_S}(x,y)$$

$$binary\ descriptor = \boxed{0101100100\ldots1}\boxed{0101\ldots1}$$

Center RF | Sign DoG

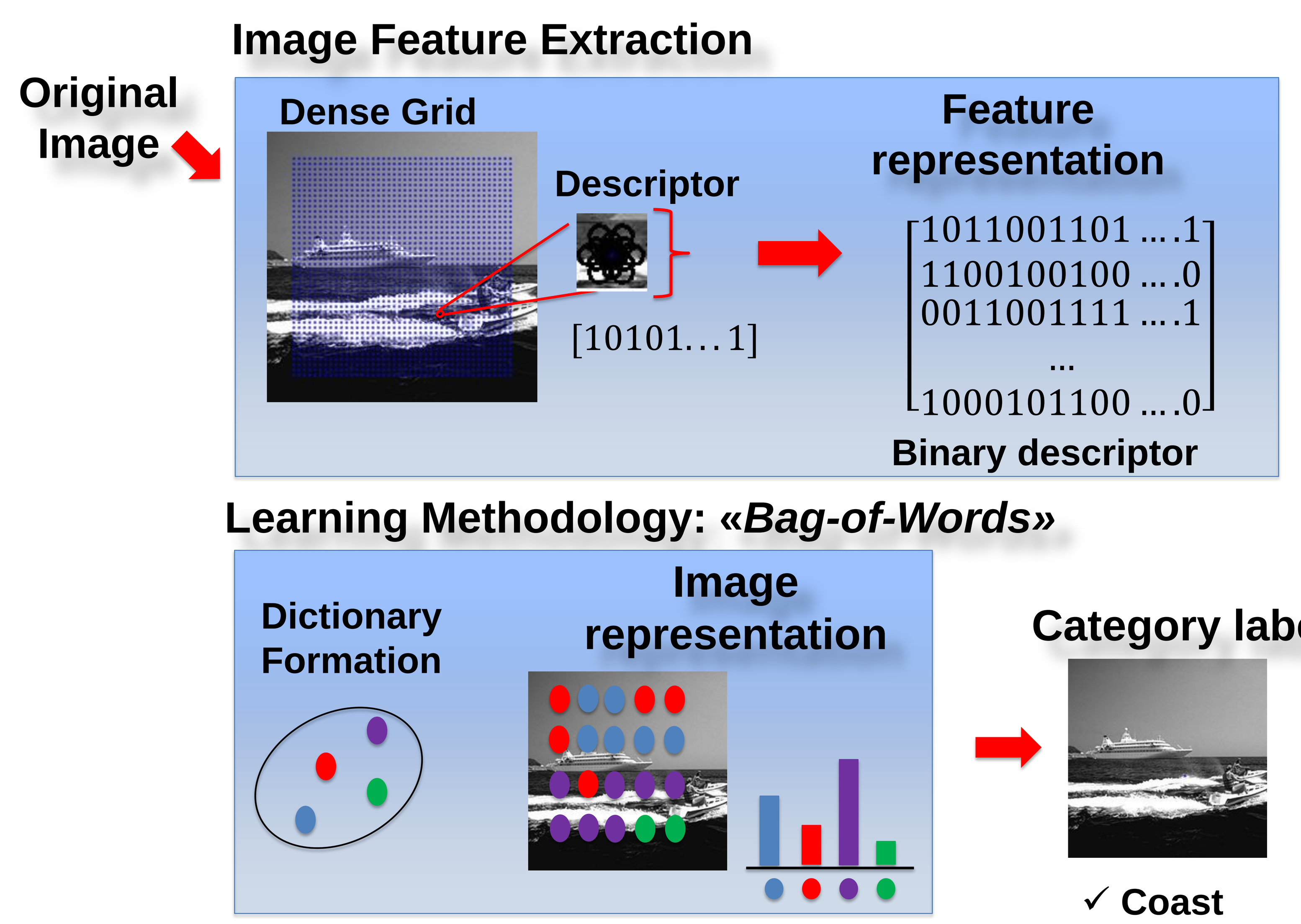
Descriptor 3: Enrich the descriptor adding ON and OFF cell responses

$$A_{+1} = f(+(I * K(x,y))) \qquad A_{-1} = f(-(I * K(x,y)))$$

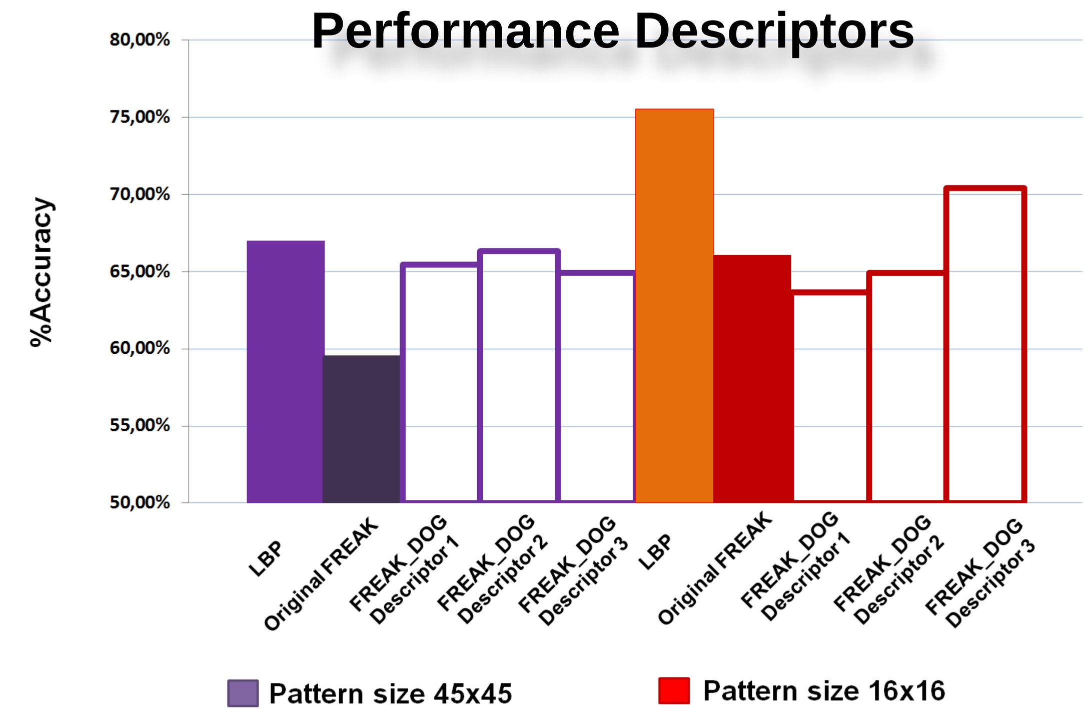
$$binary\ descriptor = \boxed{0101100100\ldots1}\boxed{0101\ldots1}\boxed{0101100100\ldots1}\boxed{0101100100\ldots1}$$

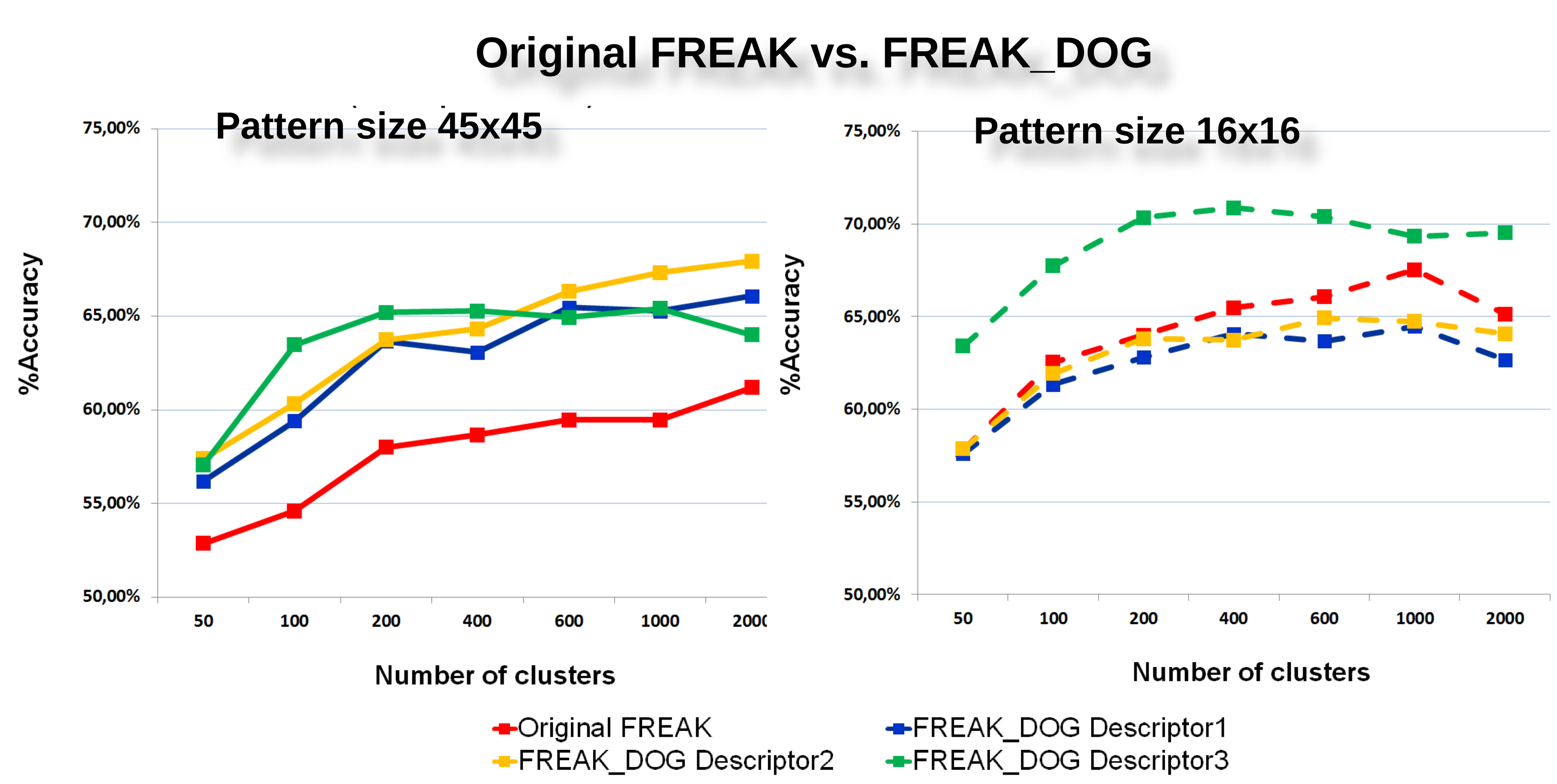
Center RF | Sign DoG | ON cells | OFF cells

3. Pipeline



4. Results





- Our preliminary results shows that **enriching FREAK** with **ON and OFF RFs** improves the performance of the original FREAK.
- In order to evaluate the impact of retinal geometry, our model had been compared to **LBP**. From our data, it seems that the **size of the pattern** makes a difference in the results, more than the organization of the RFs.
- In the future, the model will be enriched considering **other processing performed by the retina**.

References

1. Alahi A., Ortiz R. and Vandergheynst P. **FREAK: Fast retina keypoint.** CVPR 2012, pages 510-517.
2. Croner L.J.and Kaplan E. **Receptive fields of P and M ganglion cells across the primate retina**. Vision Research, 35(1):7-24, 1995

Acknowledgements

This research received financial support from the 7th Framework Programme for Research of the European Commission, under **Grant agreement n° 600847: RENVISION** project of the FET programme-NBIS

RENVISION