Aim: Study the sparse representation of natural image statistics as the receptive fields of simple cells in primary visual cortex

Several computational studies conducted since Barlow made this proposal have demonstrated more concretely the relationship between sparsity and the statistics of natural scenes. In the visual system, for example, the images that fall upon the retina when viewing the natural world have a relatively regular statistical structure, which arises from the contiguous structure of objects and surfaces in the environment. Field has shown that the receptive field properties of simple-cells in primary visual cortex (V1) are well suited to this structure, in that they produce sparse representations. Olshausen had subsequently showed that, when the receptive fields of an entire population of neurons are optimized to produce sparse representations, that the set of receptive fields that emerge resemble those of simple-cells.
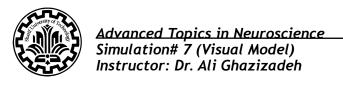
Exercise tasks:

1. **Simulate sparse basis functions of the natural images:** Using the IMAGES.mat file, which contains 10 natural images, and the method proposed by Olshausen, simulate the basis functions for sparse representation of natural images. Actually, you should simulate part "a" of the figure 4 in the paper "Olshausen, Field 1996".
   In order to solve the conjugate gradient descent part of the algorithm, you could use any preferred package or code.

2. **Study the effect of different datasets**: here we try to find the basis functions for different datasets and compare them with the basis drawn from the natural images.
   - From the following link:
     http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/ExtYaleB.html
     download images from the Yale face dataset and do the previous section and find and show the basis functions.
   - From the following link:
     http://yann.lecun.com/exdb/mnist/
     download the MNIST dataset and show the basis functions for the hand written digits dataset.
   - Show the basis functions for a different natural image dataset like Caltech101

     Some notes:
   - Randomly select 10 images from each dataset and do the sparse representation only for this 10 images.
   - Images have to be whitened before entering the sparse representation phase. Some information about whitening are provided in the supplementary file.

3. **Study the dynamics of the sparse coefficients**: find the sparse coefficients for the BIRD video in the attachment. Select each 10 frames as one patch for finding the basis function and the sparse representations. How the sparse coefficients change across time?

Optional:
Here we would like to study the role of attention models in the basis functions extracted from the sparse representation method.

According to the previous session of the course (visual search and attention), we studies some computational models for attention which extract the saliency maps of the images based on the low level features.

To find the basis functions for sparse representation, we used random patches selected from the entire images. But, here we would like to confine our self to the patches extracted from the salient parts of the image. To do so, use the GBVS algorithm for saliency detection in the image dataset. Then, extract the patches from the salinet parts of the image.

Is there any difference between the basis functions extracted from the salient parts of the images and those extracted from the whole image?