Advance Neuroscience HW9

Ali Ghavampour - 97102293

1 About Natural Images

Natural images have particular statistical features. The distribution is very different from white noise or randomly created images. Natural image distribution is a Gaussian around zero. In other words, relative to the mean luminance, there are many more dark pixels than light pixels. Also, we can naturally distinguish images of the natural world from unnatural images.

2 Sparse Representaion

2.1 • Paper Images

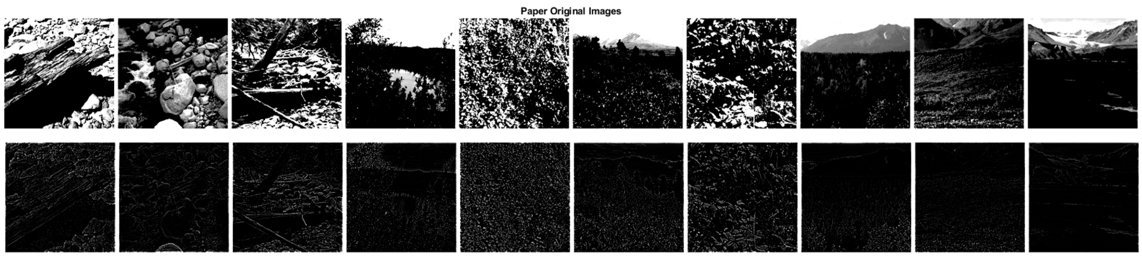

Figure 1: Top: Original Images , Bottom: Whitened Images

The sparse representation is shown in figure 12.

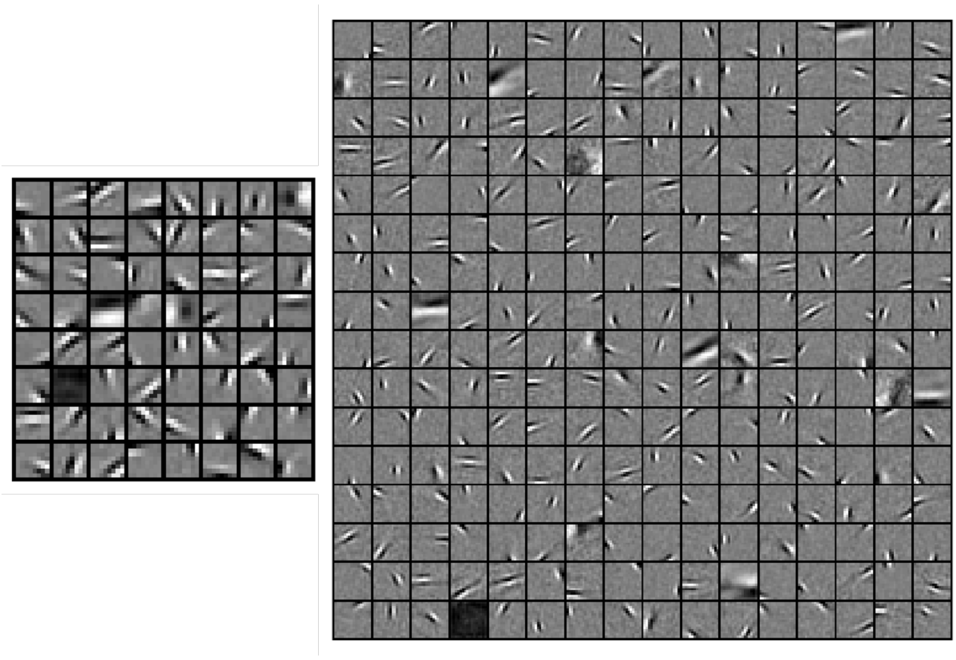

Figure 2: Sparse Representation - Left: 64 basis , Right: 256 basis

2.2 • Caltech101

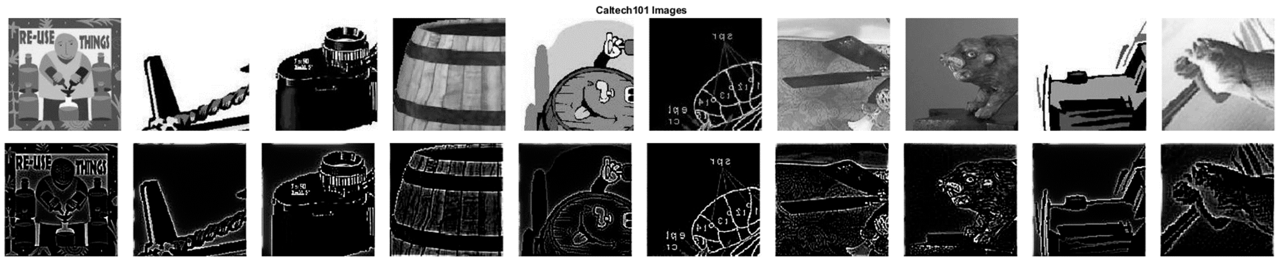

Figure 3: Top: Original Images , Bottom: Whitened Images

We can see shapes similar to Gabor in the sparse representation in both 64 and 256 basis function models. But the Gabors are smaller and more point like in comparison with paper images. So maybe we can conclude that this representation is sparser than the one for paper images.

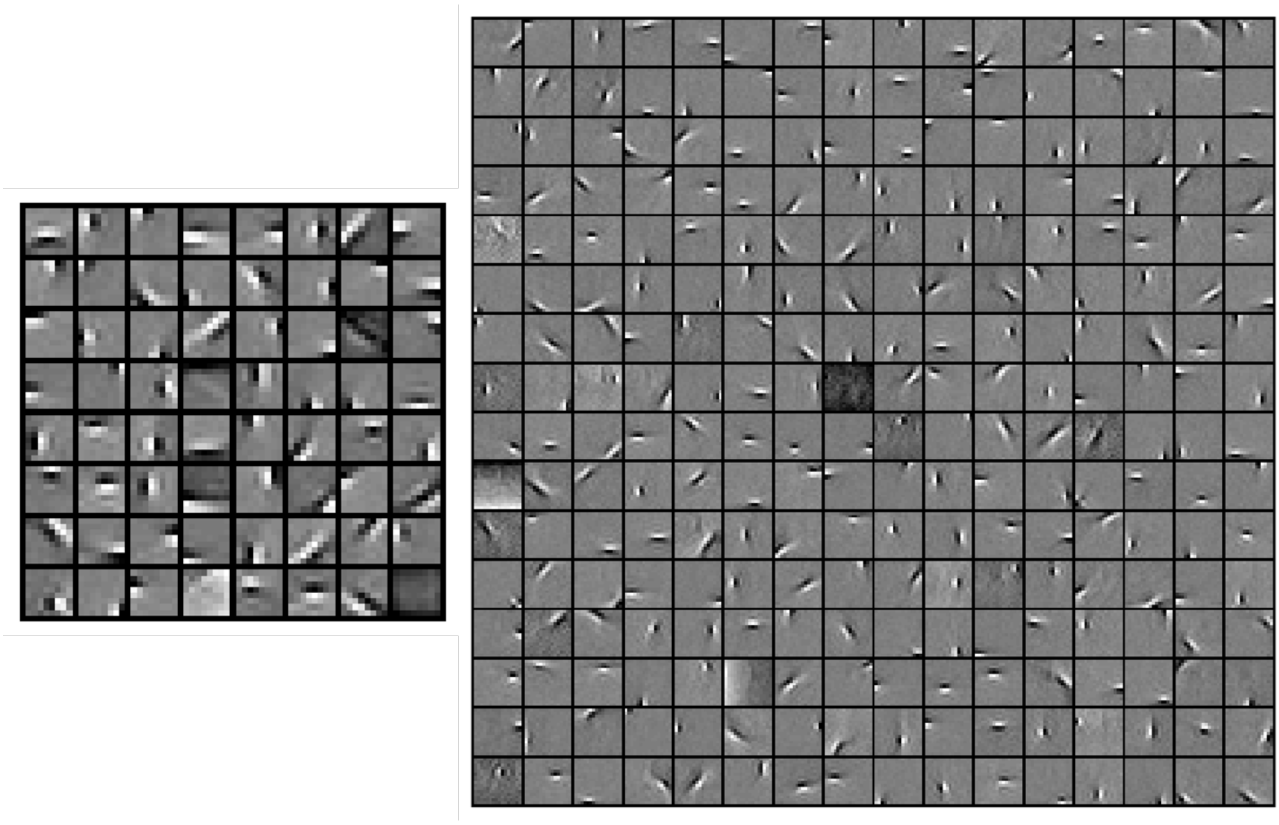

Figure 4: Sparse Representation - Left: 64 basis , Right: 256 basis

2.3 • Yale

Yale database consist of images of human faces in a dark and wague background. The images are whitened using algorithm that is provided by the writers.

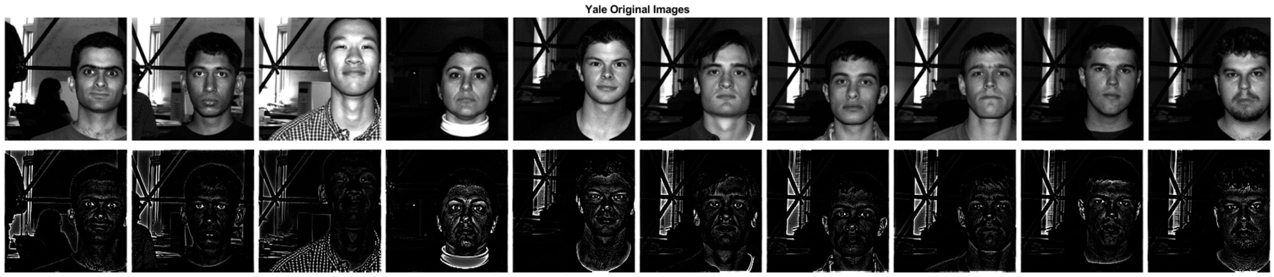

Figure 5: Top: Original Images , Bottom: Whitened Images

Here in 64 basis function model, we can barely see some orientations but the basis functions are not sparse. In 256 bsis function model, we can see more Gabor like shapes but still it is not actually sparse.

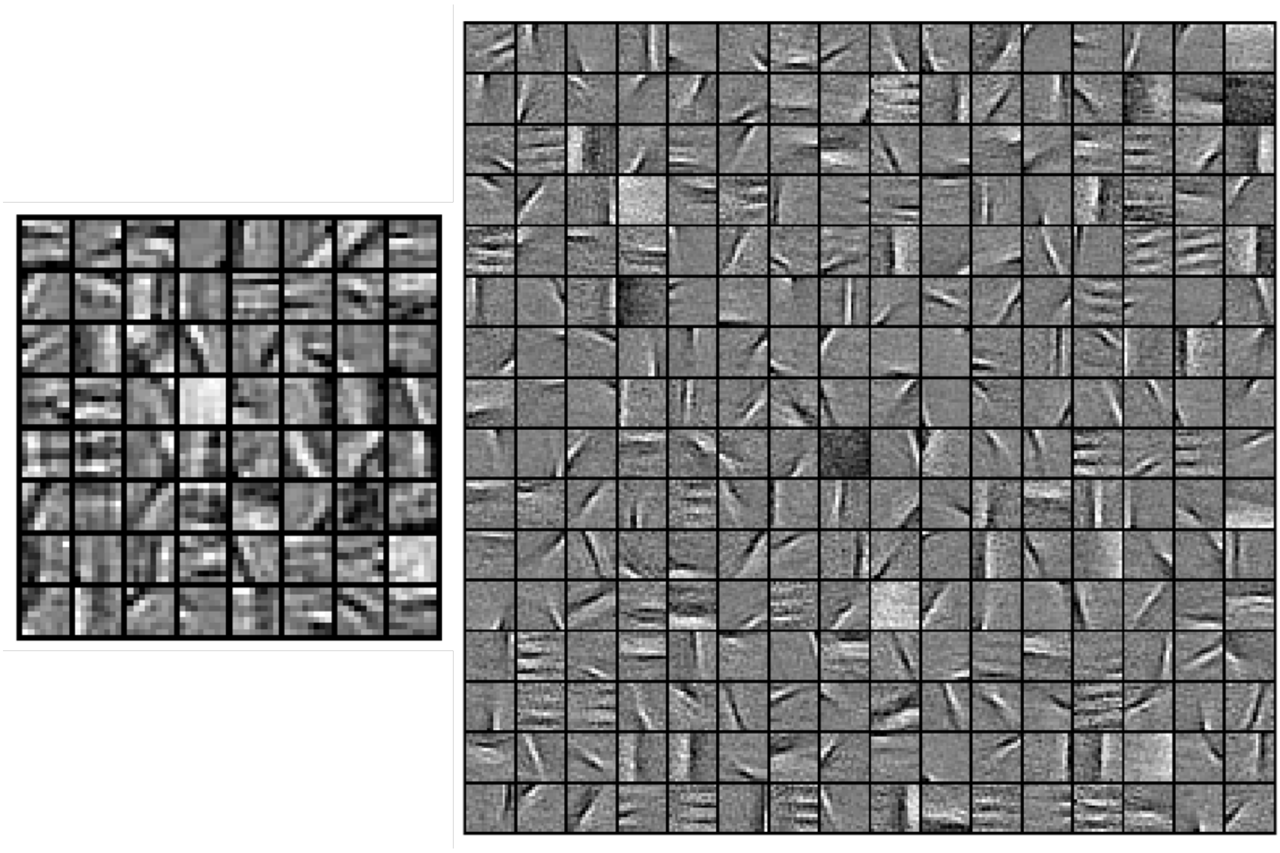

Figure 6: Sparse Representation - Left: 64 basis , Right: 256 basis

2.4 • Mnist

Mnist database consist of handwritten digits and letters. The images are not natural in this database. Note that the images are whitened using algorithm that is provided by the writers.

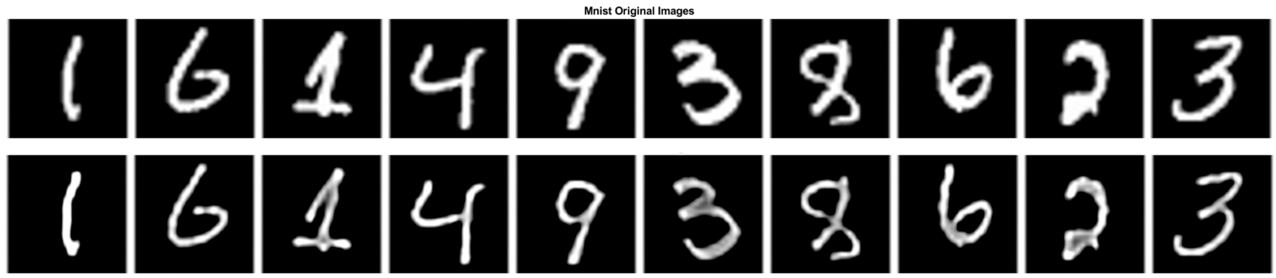

Figure 7: Top: Original Images , Bottom: Whitened Images

The 64 basis representation is the most sparse among others. But the other two are not sparse at all (especially the 256 basis one). So we can see that sparse representation

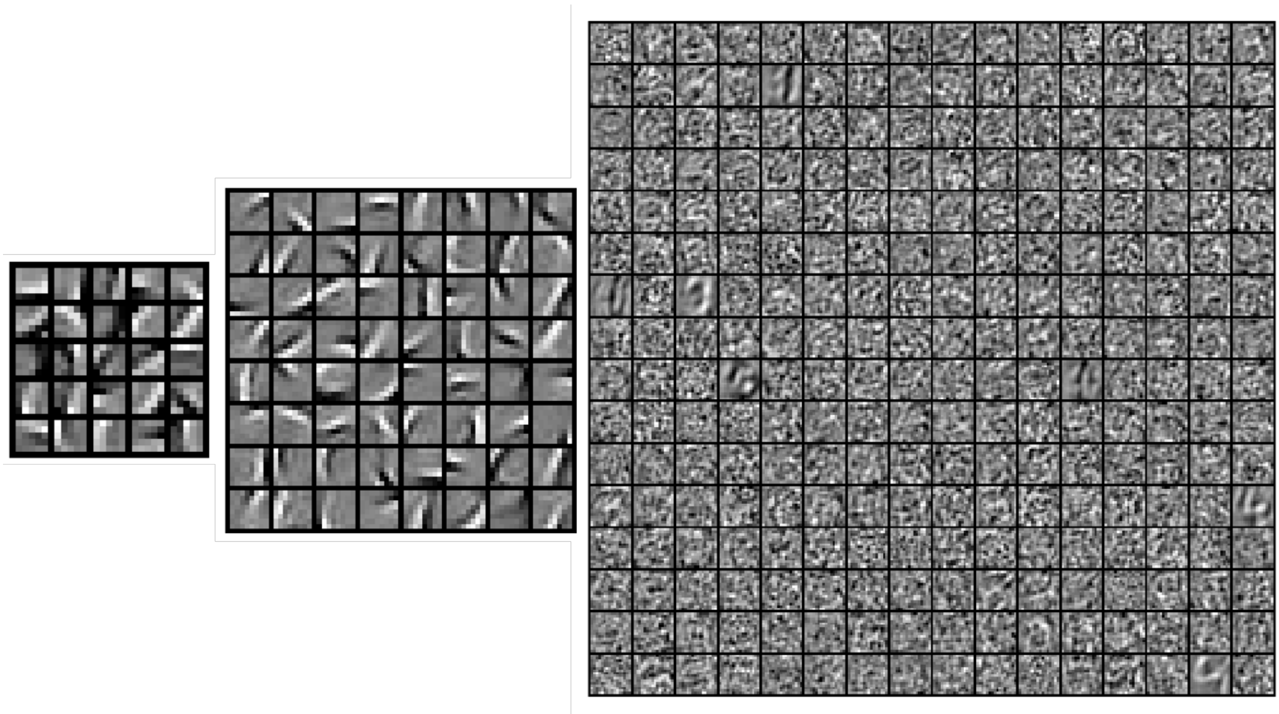

Figure 8: Sparse Representation - Left: 25 basis, Middle: 64 basis, Right: 256 basis

2.5 • Selected Unnatural Images - Set 1

These images are all made only by AI.

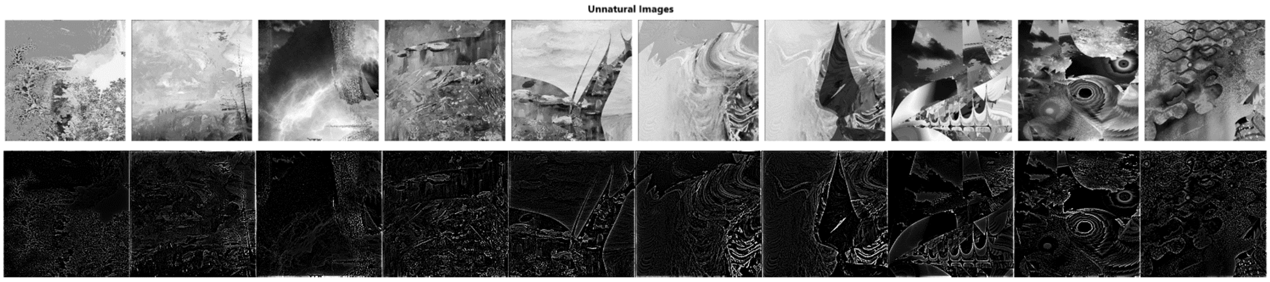

Figure 9: Top: Original Images , Bottom: Whitened Images

We can see that the representation are sparse and Gabor like although the images are unnatural. This may be because of the kind of images. The images are probably partially natural images so it becomes sparse.

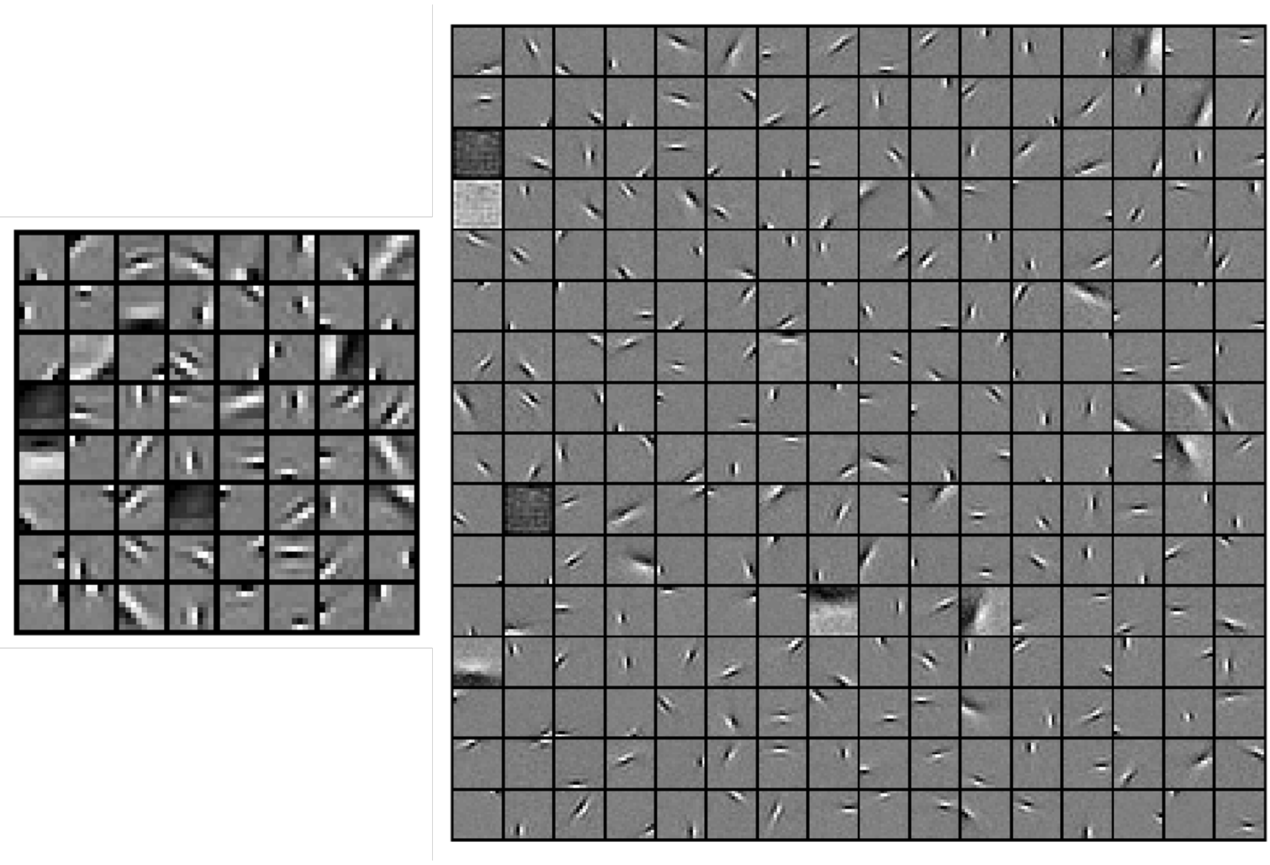

Figure 10: Sparse Representation - Left: 64 basis, Right: 256 basis

2.6 • Selected Unnatural Images - Set 2

These images are made by me in paint and some of them are screenshots from windows screens. Two of them are also random noisy patterns.

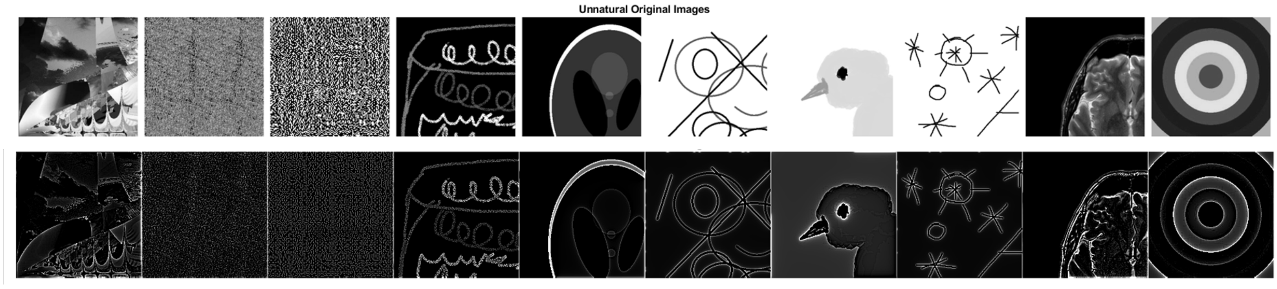

Figure 11: Top: Original Images , Bottom: Whitened Images

We can see that the basis functions are less sparse. Also, there are less Gabor like shapes.

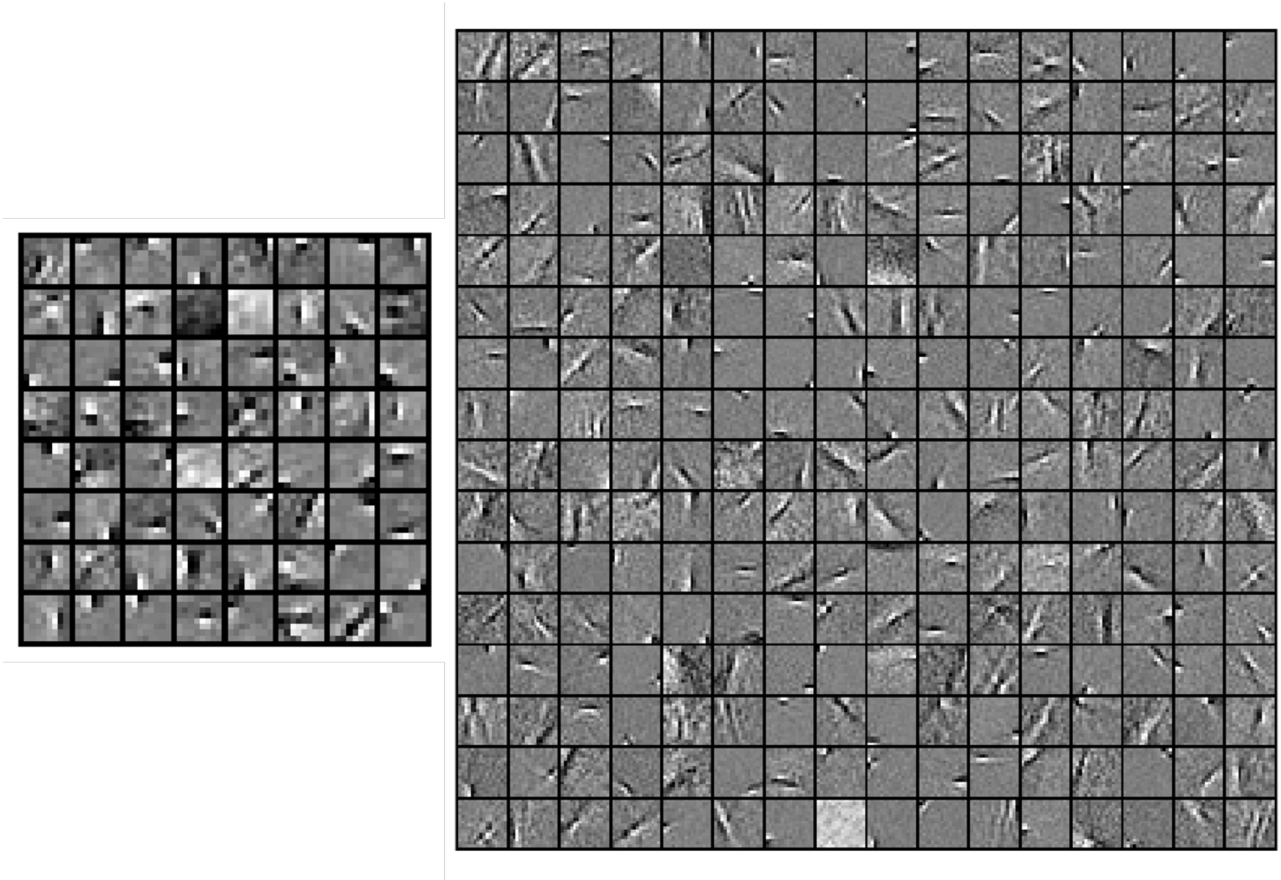

Figure 12: Sparse Representation - Left: 64 basis, Right: 256 basis

3 Coefficients in Time (Bird Video)

In this question, every 10 frames of the video is extracted from the video and is whitened. There is a total of 12 frame sets (the last set has 8 frames). Then a set of 64 basis function is learned for 10 first frames. Then the images are divided to the patches of size 8 (1296 patches per frame) and the coefficients are found for every frame set. Figure 13 is the histogram of coefficients for 12 frame sets. We can not see any important difference between histograms.

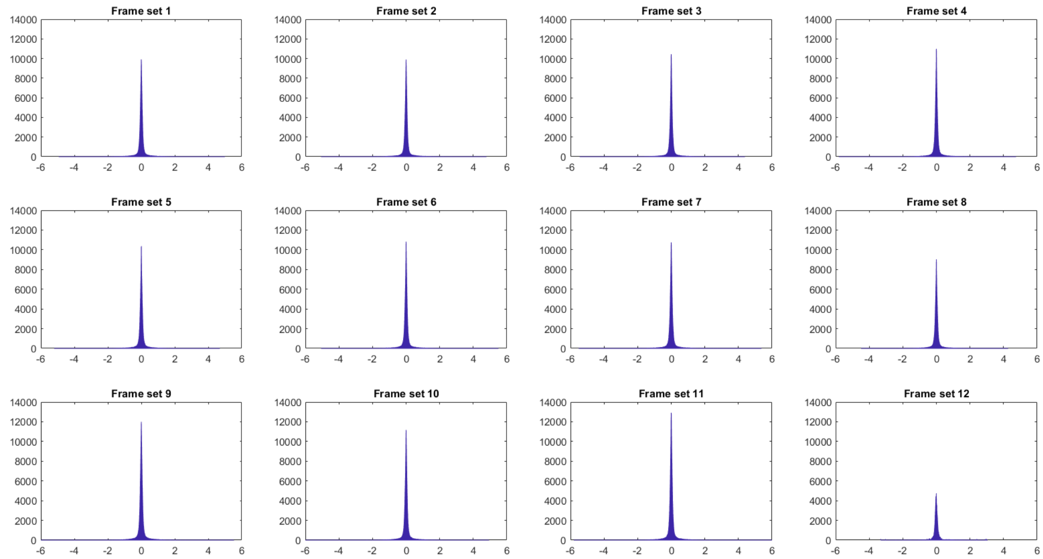

Figure 13: Coefficient histograms in time

In order to be able to compare coefficients in time, we plot the absolute of the coefficient for 4 of the patches in the 12 frame sets (figure 14). We can see that first, the coefficients are sparse (that we expected because the images are natural). Also, we can see that the absolute of coefficients are very similar to each other in different time. Especially some of them.

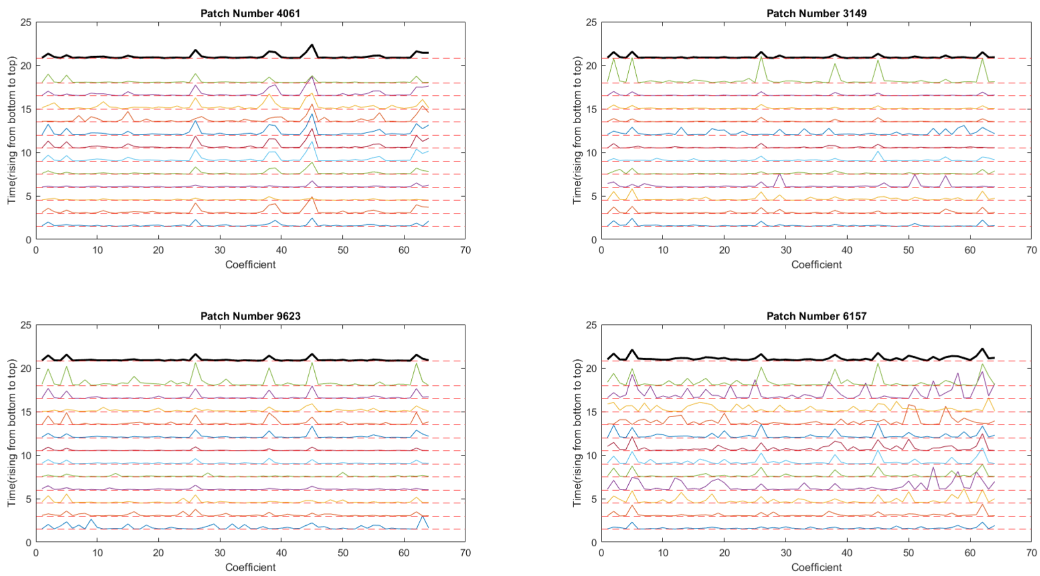

Figure 14: Coefficient histograms in time

Because of this similarity in different patches and different times, we suspect that if we take the average over all the patches we may get a similar pattern in different times. In figure 15 we can see that this suspection is true and the coeeficient are not changing much in time.

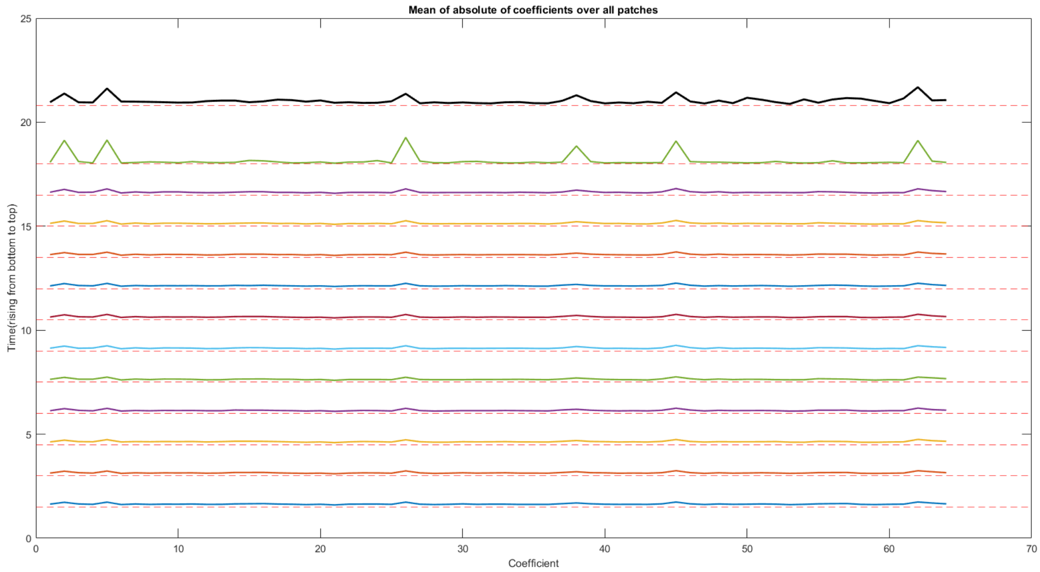

Figure 15: Coefficient histograms in time

This could be because of the fact that these maps are built based on V1 area. V1 area does not encode time effects and this code tries to reconstruct what is happening in V1 area. So, since the images do not change a lot in time, we will not see any difference.

4 Optional Part - Sparse representation for salient patches

Saliency maps are calculated for the paper images. The the images are patched into patches of size 8 and overlap of 50%. Then for each patch a saliency score is calculated. The score equals to the sum of intensities in the corresponding patch in the saliency map. Then, for different score thresholds, the patches are selected and then inserted into sparsenet.
In figure 16 we can see that for low thresholds, the basis function are sparse and Gabor like just like the one in the first question. The salient areas are the places of edges and high frequency areas in these images. So we expect the basis functions to be learned faster and also become more sparse for the patches from salient areas. But for very large thresholds, there are not enough patches for the network so it will not learn good basis functions. (As can be seen in bottom right plot of figure 16)

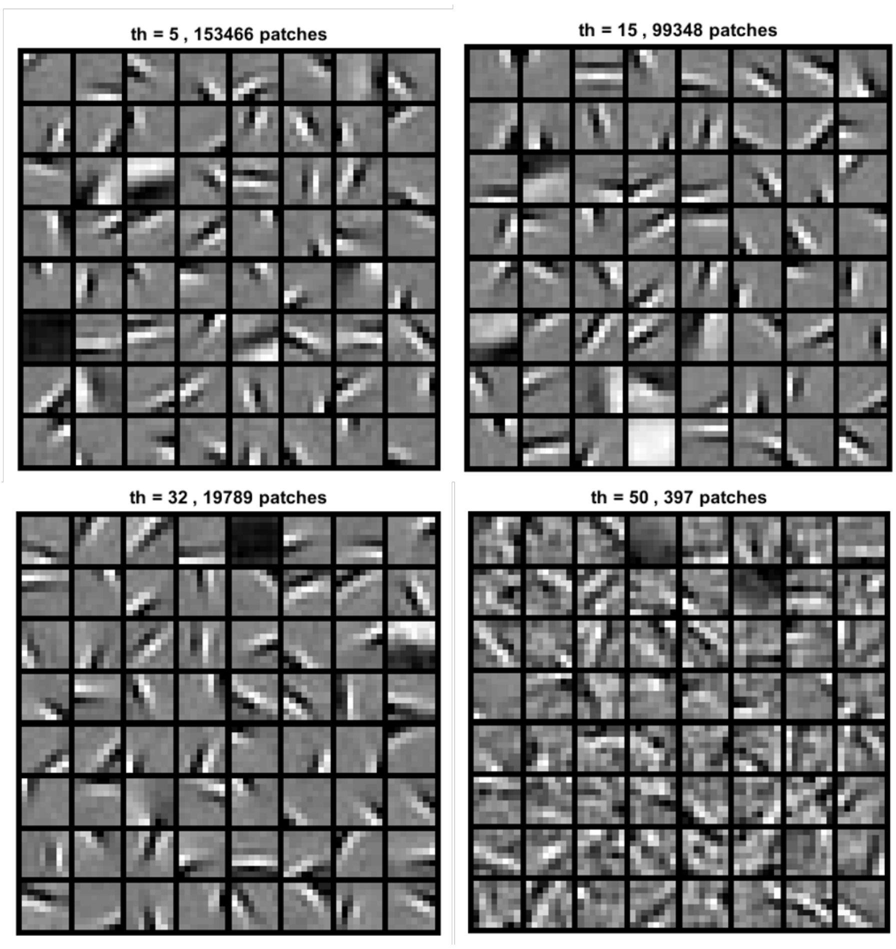

Figure 16

Also we learn the basis functions based on the patches from non-sailient areas. We can see that in those areas the learned basis functions are not Gabor like and are not sparse. So, again we can observe the effect of edges and orientations in the salient areas. Actually there is no special data in the non-salient areas so the functions are learned based on empty parts of image. (figure 17)

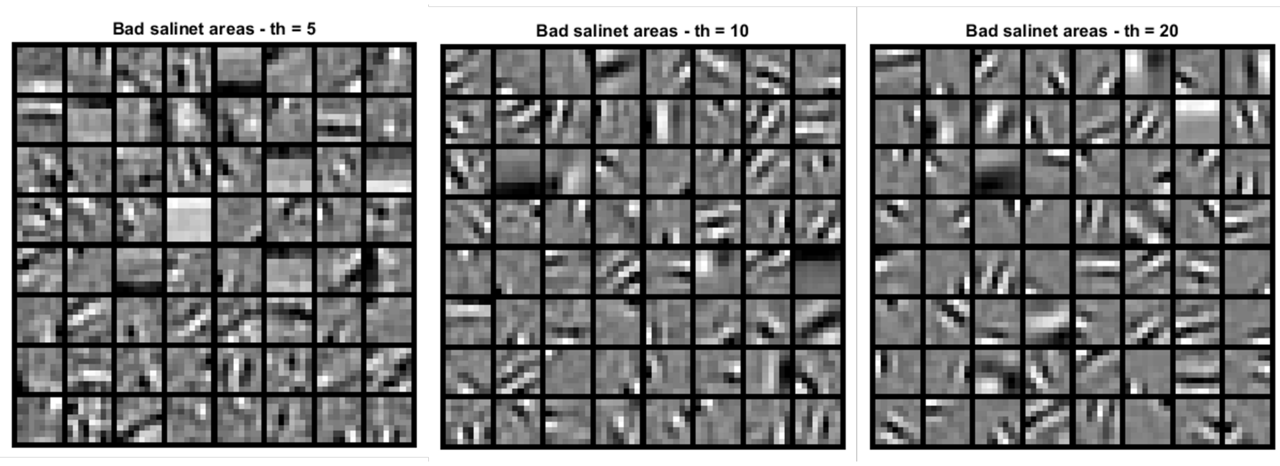

Figure 17