

# Optical Illusion

Sara Bolouki, Roger Grosse, Honglak Lee, Andrew Ng

## 1. Introduction

The goal of this project is to explain some of the illusory phenomena using sparse coding and whitening model. Instead of the sparse coding method proposed by Olshausen in [1], transient invariant sparse coding is used. Transient in variant sparse coding can learn larger bases compared to ordinary sparse coding, which can help a lot in studying illusion phenomena. Using small bases implies using small images, which in turn might diminish illusory effect that we are looking for in the image.

In this report we first briefly explain image-processing system in human brain, then the selected models for modeling retina and V1 part of the brain are discussed. Part 4 of the report explains some of the illusion effects that are going used for experiments. Part 5 contains the experiments performed and results. Finally the last part describes the future work.

## 2. Human image processing system

As we know, mammalian visual system begins from the eye. Eye can be considered as the lens of a camera that transmits the light it receives from the world objects to the retina. Retina consists of a large set of photoreceptor cells that polarize the bipolar cells. Bipolar cells in turn pass information to ganglionic cells which are the outermost layer of the retina. This information is then passed to lateral geniculate nucleus (LGN) part of the brain.

The neurons of the LGN relay the visual image to the primary visual cortex (V1) which is located at the back of the brain. Neuron responses in V1 distinguish between orientations, spatial frequencies and color in the image. The third step of information processing takes place in V2 area of the brain. V2 neurons respond to contour stimulus and are not sensitive to phase changes in the image.

## 3. Sparse coding model for V1

For modeling V1 part of the cortex, transient invariant sparse coding [2] model is used. The reason of this selection is that transient invariant sparse coding can learn larger bases, which will help us in studying illusory phenomena. In ordinary image sparse coding, each patch of images forms a column in a matrix,  $X$ . The goal is to find a matrix  $A$ , in which each column is a basis, and matrix  $S$  where each column corresponds to the activation of bases relative to each patch. Ordinary sparse coding tries to minimize the reconstructing error by achieving the highest possible sparsity. The problem can be formulated as below [2]:

$$\|X - AS\|_2^2 + \beta \|S\|_1$$

$$\text{subject to } \|a_i\|_2^2 \leq 1, \text{ for all } i$$

In translation invariant sparse coding, images, bases and activations are represented as two-dimensional arrays instead of vectors. In this algorithm, convolution and point-wise product is used instead of matrix multiplication. Applying these changes the new problem can be defined by the following formula, which tries to solve for A by minimizing the sum [2]:

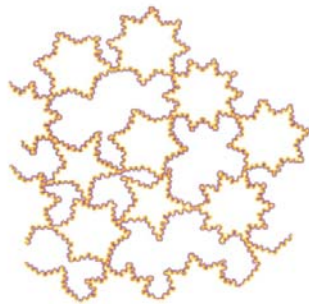
$$\sum_{i=1}^m \left\| X^{(i)} - \sum_{j=1}^n A^{(j)} * S^{(i,j)} \right\|_2^2 + \beta \sum_{i,j} \left\| S^{(i,j)} \right\|_1,$$

subject to  $\left\| A^{(j)} \right\| \leq 1$ , for all  $i,j$

The rest of the algorithm is similar to ordinary sparse coding; the first step is learning the bases and the second step is reconstruction of the image. This process is continued until bases are converged.

## 4. Optical illusion

According to literature, “An optical illusion is any illusion that deceives the human visual system into perceiving something that is not present or incorrectly perceiving what is present”[3]. Studying illusion phenomena would help us to understand the visual system in the brain. While there are lots of different types of visual illusion, this project focuses on brightness illusion, orientation illusion and watercolor illusion.



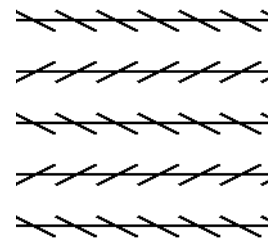
**Watercolor illusion**



**Brightness Illusion**



**Illusory contours**



**Zollner Effect**

Different types of optical illusion may correspond to characteristics of different steps in the visual system hierarchy. For instance, brightness illusion might be explained by lateral inhibition in retina and V1 receptive fields. There is also lateral interaction between neuron activities in V1 area, which may explain orientation illusions such as Zollner effect, Orbison’s and Ehrenstein illusions. The third type of illusion

that is possibly explained by these hierarchical models is illusory contours. According to neuroscience literature, illusory contours causes activation of neurons in V2 area of the brain, therefore we expect to explain them by V2 models. Since we only use V1 model in this project, we are not considering illusory contours.

## 5. Experiments

Two different experiments have been performed in this project. The first experiment tries to justify Zollner illusion and second experiment concentrates on brightness and watercolor illusion.

### 5.1. Zollner Illusion

As shown by the figure, in Zollner effect, parallel lines seem to become closer to each other. The reason of this effect might be misjudgment of the value of the angle between horizontal and diagonal lines, which could be explained by lateral inhibition between V1 neurons. The steps of our experiment are:

- 1- Learn bases from natural images
- 2- Observe bases activation for illusory images
- 3- Study bases activation considering what might causes illusion in V1 area

In order to observe illusion in the image, we need a set of bases such that the difference between their orientations is very small. Bases learned from natural images do not have this characteristic. Therefore, 360 hand made bases that have 0.5 degrees difference in the orientation are used instead.

Unfortunately, the result of the experiment could not explain the illusory phenomena, since all the activations corresponding to a horizontal line had a degree of orientation between 0 and 1. The expected result would be activation values with 3 to 4 degrees of orientation, at least.

### 5.2. Brightness and Watercolor Illusion

The goal of the second experiment is to explain brightness and watercolor illusions. Our hypothesis is that whitening is performed by the early visual stages, but brain itself does a form of inverse-whitening operation in lateral visual system in order to compensate the whitening of the early stages. The steps of this experiment are:

- 1- Learn bases from natural images
- 2- Reconstruct a group of natural images with transient invariant sparse coding using natural bases
- 3- Learn inverse whitening filter for this group of images
- 4- Reconstruct illusory images with transient invariant sparse coding, using natural bases
- 5- Apply learned inverse whitening filter to the reconstructed illusory images

#### 5.2.1. Learn inverse whitening filter

In this stage, the goal is to minimize the error between the initial image and the reconstructed image from the sparse coding, by applying a filter to the reconstructed image. The inverse whitening filter is learned in frequency domain. The objective is to minimize the following equation:

$$\min \sum_j \left\| X_j - Z_j \times g \right\|_2^2$$

Where  $X_j$  is equal to FFT of  $j$ th image in the training data set,  $Z_j$  is equal to FFT of the  $j$ th reconstructed image and  $g$  is the inverse whitening filter which we are going to learn. Assuming that there is no dependency between frequency response of each filter, we can rewrite the above formula as the following:

$$\min \sum_j \sum_s (X_{j,s} - Z_{j,s} \times g_s)^2$$

Where  $s$  corresponds to each pixel of the image.

Linear regression method is used to solve the above problem. Parameters that should be learned are different components of filter  $g$ . The closed form solution is calculated by taking the derivative according to each component of  $g$ , and setting the derivate equal to zero:

$$l(s1) = \min \sum_j (X_{j,s1} - Z_{j,s1} \times g_{s1})^2$$

$$\frac{\partial l(s1)}{\partial s1} = \sum_j -2(X_{j,s1} - Z_{j,s1} \times g_{s1})Z_{j,s1} = 0$$

$$g_{s1} = \frac{\sum_j X_{j,s1}}{\sum_j Z_{j,s1}}$$

$g_{s1}$  is the first component of the  $g$  filter. Other components of  $g$  are calculated with the same formula. 60 images with 100x100 resolution were used in order to learn the inverse filter. The learned inverse whitening filter has the same dimension as the training images. Figure 2 shows amplitude of different frequencies of the learned filter. As shown, the filter has larger values at lower frequencies, therefore amplifies lower frequencies and can be considered as the inverse of whitening filter.

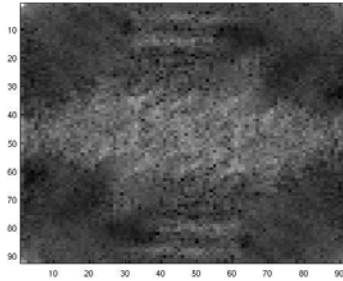


Figure 1 - Inverse whitening filter

### 5.2.2. Results

To do the experiment we needed to reconstruct the illusory images by the sparsenet code and then convert them to frequency domain. In frequency domain, we can apply the learned  $g$  filter to the reconstructed images. The next step is to apply inverse Fourier transform and bring the images back to the spatial domain. Figure 3 illustrates one example of performing these steps on a sample image in the training data set.

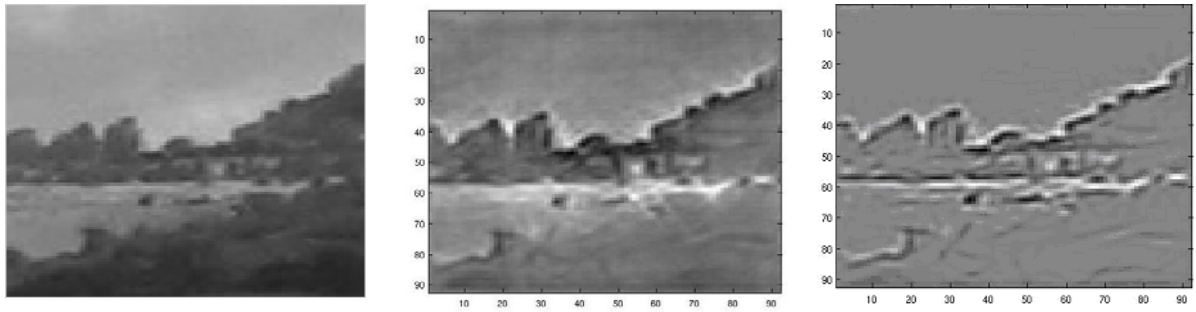


Figure-2 (left) Input gray scale image, (center) Reconstructed Image, (right) Reconstructed image after applying inverse whitening

As shown in Figure 2(left), inverse whitening tries to make the image smoother so it spreads out white and dark borders. We were expecting to be able to demonstrate brightness and watercolor illusion by this smoothing. The same test has been performed on brightness illusion and a group of watercolor illusions. Figure 3 shows brightness illusion image reconstructed by the sparse coding before and after applying sparse coding. As shown in the Figure, before applying inverse whitening we can see a lighter and a darker bar, but what we expected from inverse whitening was to make the left part of the image light with a lighter bar at the center border and the right side darker with a darker bar at the same place. As shown in Figure 3(right), the expected result is not obtained.

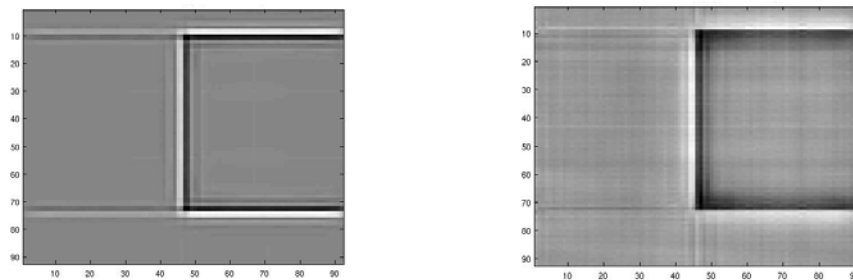


Figure 3- (left) Before applying inverse-whitening, (right) After applying inverse whitening

The same test was performed on a group of watercolor illusion images, expecting to see spreading of the inside color or detecting the illusory contours. Unfortunately the reconstructed images did not have either of these characteristics. To make the experiment more accurate the same test was performed on 200x200 images, which had the same result as 100x100 images.

## 6. Future work

There are two suggestions for future research. The first is to study other models of whitening which can also help learning and parametrizing inverse whitening filter. The second suggestion is to re-learn the bases according to the un-whitened images, meaning that the coefficients are derived based on old bases and whitened images, and then bases are re-learned for the un-whitened images and obtained coefficients.

## 7. Reference

[1] Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Bruno A.Olshausen, David J.Fild, Letters to Nature, June 1996

[2] Translation Invariant Sparse coding, Roger Gorsse, August 2006

[3] [http://en.wikipedia.org/wiki/Optical\\_illusion](http://en.wikipedia.org/wiki/Optical_illusion)