CNS 187 Neural Computation - Problem Sheet 6

Ling Li, ling@cs.caltech.edu

November 10, 2000

6.1 Readings - Using action potential timing to explain vision phenomena

I will talk about vision transient and negative image. In page 35 of the paper in Nature, it is said that "If input current levels are so high that multiple spikes are present in a single cycle, ..., spurious recognitions due to these other spikes might sometimes be made later in the cycle." However, this feature may explain why we human beings have vision transient. I know nothing about the biological aspect of vision. So below is my imagination.

1. **Vision transient**. Vision transient is a phenomenon common in our experience. When you see something moving quickly before you, the vision you get might be the trajectory formed by the moving object, as if your eyes can keep the vision for a short period.

   In the model that represents the stimulus by action potential delay, we can explain this phenomenon as: At time t, the eyes (or the retina) receive the stimulus of object at position $p(t)$. Neurons around $p(t)$ will firing at some delay, and make the followed coincidence detection neurons believe some object is there. At time $t + \Delta t$, the image of the object moves to position $p(t + \Delta t)$, the neurons around $p(t + \Delta t)$ begins to fire. However, neurons around $p(t)$ may still fire, which is called 'multiple spikes' in the paper. Refer to Fig. 2 in the paper, the relative time delay among spikes from different neurons around $p(t)$ will remain the same, or roughly the same. This explains why we can still 'see' that object in $p(t)$ in a short time after t. Further, we can imagine that with the decreasing input current at $p(t)$, the relative time delay will change with the time, and after some time the spikes will disappear due to the more and more weaker stimulus. This explains why we will not always see that object, and why the vision transient is vaguer and vaguer with time.

   Similar phenomenon happens with peoples' ears. After some strong sound, you will still hear that sound for a while.

2. **Negative image**. This is also a common experience. When you stare at something for a while and then close your eyes, you will still see that object for quite a while, with a negative image. I think this is also a kind of vision transient. However, why negative image?

   I guess that the eyes receive two kinds of stimuli: one is the brightness of the object, the other is the color information of the object, which may further consist of several primitive color features, such as red, green and blue. I argue that neurons will increase their thresholds or decrease the sub-threshold activity while continuous strong stimulus is received. When a red object is stared for a long time, the thresholds for red stimulus are raised or the sub-threshold activities become weaker. Then when you close your eyes, no red spikes exist while the spikes for other neuron, though very little, becomes dominant. And, the multiple spikes of the bright-sensitive neurons will still exist for some while. Thus the brain will compensate the bright signal with colors other than red. Thus the negative image produced.

6.2 Final Project Initial Report

First, I want to list something that seems to be weird.

- Animal 1 Cell #63, which is in area A, acts different to actual same input. For experiments #2–5 with 213Hz tone, the number and the start position of spikes varies largely. However, they should be similar according to the role of neurons in area A. For example, in experiment #2, after a large delay does the neuron start to spike. It is not coincide with the experiment #3, which has the same input in the beginning part, and has a much earlier spike.
- Animal 1 Cell #1, a γ cell in area W, takes the `S000_176.wav` which is a 20-tone imitation of a good spoken "one", as "one". If standing at a human's position, that is not a voice.
- Male voices are usually not accepted by *mus silicium*, though it is announced to be.

OK, I do not want to say that *mus silicium* itself is weird. These looked-strange response must have some inner reasons in the *mus silicium*. That's why I collect some of them here. I think several principles below is adopted by the *mus silicium*.

- Type γ cell use only frequency information from α- and β-type cells, and the frequencies it uses are not continuous. More specifically:

  As mentioned in the paper of the *mus silicium*, the cells in area A are sensitive to frequency and are not so sensitive to the power of the tone (Fig. 4c in paper). Thus, only frequency information is inputted to the layer W.

  Further, the cells in area A are declared to transform the sound file into 40 channels of stimuli, each channel corresponding to some specific range of frequency. Frequencies in the same range will be regarded as the same.

  Thus, by only changing strength of some frequency or moving the frequency a little, the *mus silicium* should still think it is "one". This is what `S000_176.wav` told us. We may also check the file `S000_177.wav`, which is refused by the *mus silicium*. I think the reason is that it has not enough frequencies. (However, no other similar sound files found, and the server is really too busy to accept new experiments. Otherwise we can dig more.)
- As described above, the generalization of the γ cell relies on how frequency information is retrieved from the voice signal, and, on how frequency information is used by α- and β-type cells. The processing of the frequency information by α- and β-type cells, as we can imagine, is to extract specific combinations[1] of frequencies. The γ cell then uses the sequence of recognized combinations to recognize words. Thus we can expect the generalization in what time you begin to say "one", in what strength, and with some frequency floating. However,

[1] Or sequence? I think it should be combination.

we can not expect too much on the generalization on the changing of frequency structure of the voice, such as male v.s. female, increasing tone v.s. decreasing tone,[2] since the γ cell was trained by a single woman.

I am thinking whether adding another one-layer conventional forward-feed network after the layer 2-3 will help.

Several things to be dig out:

- How to train the *mus silicium*? How to train with several different samples (may vary much in time sequence and thus may vary much in the time delay)?
- How α- and β-type cells recognize frequency combinations? Winner-take-all?

[2] I am not so sure about the difference between increasing tone and decreasing tone. Some sound files with different tone to the trainer are recognized, while some not. Maybe there is some other reason.