# Computational Neuroscience

**Dr. Sharba Bandyopadhyay**
**Department of Electronics and Electrical Communication Engineering**
**Indian Institute of Technology Kharagpur**

**Week – 12**
**Lecture – 56**

Lecture 56 : Optimal Coding in Visual System

Welcome. So we are into the last week of our computational neuroscience course. So far we have covered number of topics starting with how spiking happens in single neurons, how it propagates along axons and then at the synaptic terminal neurotransmitters are produced and its information is transferred on to the next neuron through current injection at the postsynaptic side. And we have discussed how the neuron producing action potentials, how the series of action potentials produced due to stimuli can be understood and studied to understand what kind of features of the stimulus the neuron is encoding. We also then carried on to understand how from the spike trends we can decipher what stimulus was there and so on. And finally we had our discussions on plasticity, short term and long term, where we saw how neurons adapt because of certain, through activity, input and output activity and self-organize themselves.

We saw a few of the learning rules. We also saw how these rules can be implemented in model neurons. So we also looked at certain phenomena that can be explained with plasticity. And so this was overall sort of the whole structure of the course and I hope that we got you across the ideas behind all these topics that we had discussed, the primary ideas.

But we also want you to know that so far what we have talked about has been based on actual experimental data. So we have been seeing this from the point of view of an empirical science in the sense that we have some sort of a hypothesis and we test that by designing certain stimuli or designing a certain type of experiment and we then go on forward to record activity of neurons. And then based on those activities we make conclusions about the process. And we have also talked about how manipulation of that activity also is required in order to understand causal rules. So with this thought, I mean obviously the questions that arise is that, is it that we will have to keep on doing such experiments, designing such experiments and then performing such experiments in model organisms or in animals that are representative of certain kinds of functions or certain kinds of disorders.

And then we test those ideas through those experiments and the data that we collect. There is of course a limitation in terms of performing experiments. We obviously cannot, at least of today, we cannot record activity of all possible neurons that are involved in a particular phenomena. For example, even if we talk of coding of a single sort of sound token, we are unable to record activity throughout the auditory pathway and beyond. And even in the cross model active structures that are also slightly activated, maybe even in the sub-threshold level in order to understand exactly what is going on.

So obviously we need to make sort of extrapolations based on whatever results we have to other kinds of stimuli. And maybe frame hypothesis, but ultimately these have to be tested out. So our results may indicate some prediction, but unless we have tested it out, we are never sure because we do not have any laws that are followed. At least of now, we do not know of any laws that can be used to sort of at the overall glow level for the neural systems that can be used to predict behavior of certain neural system for a certain kind of input or in a certain kind of context. So we obviously now need to also think about ways.

Well, I am not saying that we should not be doing experiments because everything that we know is based on experiments, such kind of experiments. And so without experiments, there is no way forward. However, we should start to think, I mean, I would like you to start thinking in these lines that how can we approach the problem from more sort of fundamental principles. Of course, there are no laws, but are there things that we can use in order to study the neural system through theoretical means, through analytical means. And then make predictions about the system and see how the system should be under certain kind of constraints.

So with this background, we will start our discussions on some theoretical work on coding of sensory stimuli. So we will talk about the visual system, one particular work that was done by Olsouw and Andfield back in 1996. And there have been some extensions of that work, but it has not been followed up extensively to more detailed things and more higher levels in terms of higher levels. So let us think of what are sort of the principles that may be useful. So if you think of the brain, we know that we have action potentials, that action potential is encoding particular type of information in some transformed way of some phenomena that is producing that action potential.

It may be a stimulus or something in the external world that forms those action potentials in the entire network in the pathway and at every stage it forms some kind of information or some part of the stimulus information is present in those action potentials that is carried forward. And finally, through the population activity of neurons, we perceive that stimulus and based on that we guide our motor system to produce a behavior. If we just think of this sort of loop, I mean the behavior can then again produce another stimulus but either an open loop or with a feedback completed loop, we can think of it in both ways. So what are the sort of constraints in this whole idea? That we need to have complete information about the stimulus in those action potentials. When we mean complete information, the things that we will discuss is based on complete information about the stimulus physical parameters.

But that is not necessarily the complete information. Of course, if we have all the physical parameters of the stimulus, we have the complete information. But it may be way more than what we actually require in terms of what needs to be encoded. Because there are, I mean, if you think of speech sound, vowel sound can be heard based on simply tones present at the fundamental and its formant frequencies. And we hear the vowel sound almost equally well.

I mean you can create tonal speech like that. And so there is much less amount of parameters that are actually required to form that percept. So this was an aside in the sense that we right now really do not know how to formulate the problem based on that kind of information in the sense that only what is relevant in the physical parameter space of the stimulus. For our discussions here, we will limit ourselves for information in terms of simply the entire set of physical parameters for the stimulus. That is if we think of a sound, it is the entire sound waveform.

But we actually, I mean, in the ideal sense, we do not need the entire waveform to understand the sound. Or the brain does not actually encode that entire sound structure, the physical sound structure if you will. It is probably in some other transform space that information is present in the action potentials. But since the whole, if we have the whole information, the other kind of information is also there, so we are not losing that out. So the one thing is that we need to keep the information intact in whatever sense we mean information so that we can get the percept of the stimulus.

And in order to frame, I mean, if we think of up to this point, we will not talk about the motor and behavior part, up to simply the feed forward thing, up to this point to represent that information, what we are using is action potentials. So action potentials is activity of neurons and if you remember, it is what is reflected in terms of how much energy is being used up in regions of the brain when we do functional magnetic resonance imaging, where the amount of oxygen consumed in that particular region reflects the amount of activity in that region. And just from, what should I say, just from common sense in the way that we want to minimally utilize resources, we want to reduce the amount of resources that we want to use because that way, if we think of it that the system has organized itself through evolution in some way, it probably has done some optimization in the process and in that optimization, it is very likely that the amount of energy consumed for a particular task is minimal. There is no rule like this or there is no proof of this or anything, I mean, this is almost like an axiom if you will that, okay, minimum energy use is there, is something that we should be following or nature has followed in the process of evolution in developing these systems. So in other words, we basically have two things here.

One is energy usage and information being intact. Information loss is minimized. So energy usage is minimized, information loss is minimized. So if we can formulate the problems of representation of general stimuli, let us say for the visual system, then can we get any idea about what the system should be like or should have been like at least at one level. This is what the Othars, Olsouwsen and Field approach, this is the question that the Othars, Olsouwsen and Field approach that, okay, so energy usage is minimum, which means that if we have, if we represent functional units in the neural system at at least one level, let us say, then their activity or the spiking or their rate, which represents the amount of energy that the system is using, that needs to be minimized and yet the loss of information is also minimal.

So both of these things would be minimized and when we mean by loss of information, what we will mean in this case is that based on the spike trends, we should be able to reconstruct the original images, which again, as I had mentioned earlier, need not be the case, but for this starting point, this is what we will be going with. So for energy minima or minimum utilization of energy, what they use the idea of sparse representation. Sparse representation is essentially an indirect or way of saying that the energy usage is minimized or reduced. So what is represented here, so the work they did was on natural images. So if we have an image that is described here, that rectangle, I mean, the parallelogram that you are seeing, that is being represented by the activity of many units or neurons and AI is representing the activity of the i-th unit that is representing the image.

So what is being depicted here in the dots here is that only a few of those units are active with high amount of activity, while most of them are not active or minimally active. So and this is for different images, different sets of units will be active with higher levels, while others would remain near zero values. So this idea of sparseness was incorporated by them by using distribution of the activity to be of this nature which has huge amount of mass near zero. So if we have a system that has let us say capital N neurons and they have activities A1, A2 up to An for stimulus 1 which may be given by 1, 1, 1 here for stimulus 2 or image 2 it is A21, A22 and so on A2n and like that we have capital M images AM1, AM2 up to AMn. So what they are saying is that if we take any one unit AI, let us say the i-th unit, i-th row here, so for all the capital M images the activity of AI would have a distribution that is, that's mass is concentrated near zero that is most of the time that unit is inactive whereas it has a heavy tail in the sense that it has lower and lower probability of having a high activity.

But it can have higher and higher activity and only in very few cases, fewer and fewer cases. So this would be the constraint that they are using in terms of energy usage and in terms of information not being lost what they did was use the image reconstruction error. So what they have is they are representing, so if you take an image, let us say this i, we can create based on the pixel values this i can be represented by let us say a vector i x where each of the pixels are then expanded out in the form of a vector. This i x is representing the entire image where x is the image vector. Now we have, they are using the basis of the activity of the units by phi i.

So that is akin to the receptive field of the i-th unit. So phi i operating on that vector length x is going to produce and multiplied by the activity level a i is what information it is keeping about the image. So $a_i \phi_i x$, this term together with all the units i equals from 1 to capital N as we saw, I am sorry capital N as we saw, these capital N units now have these activity a i's a 1 to a capital N and for one particular image the difference between the two is the error in every pixel and the mean squared error is represented by squaring that term and summing over x and they additionally have the term s a i which is a penalization of incorrect sparseness. That is if a i is large, too large, it is penalized by a weighting function that is represented here. So that is if a i is low, then there is no penalization or very low penalization and as if a i is large, then it has to be penalized with higher amount.

So and lambda is simply a factor that weighs the two things, two quantities that we are minimizing and providing the relative sort of weightage of the two. So this is providing the error term and this is providing the sparseness term and now this overall function is minimized so that the error is minimized and the penalization to departure of sparseness is reduced. So by doing this and then you can use standard optimization approaches, you can use different kinds of functions here for the use log 1 plus x square. In this particular case, there are other functions that can be used. Of course, there are factors that are required here which can be taken care of by lambda.

But the basic idea is that now you do a gradient descent based on a huge set of natural images and gradually see what kind of phi i's which is representing the receptive field of the neuron, what kind of phi i's emerge from this kind of optimization. That is what sort of receptive field should we need. This is what was happening that the energy needs to be minimized and the error needs to be minimized. So what they found with this and some other methods with the same ideas is that they found receptive fields of these so-called theoretical units to be very much like simple cell receptive fields about which you learned a little bit in the beginning towards the beginning of the course. So what is being represented here are all the phi i's.

So I think there are 144 12 by 12 phi i's, phi i meaning the 144 receptive fields which are used to operate on the image at every region and then each of these based on the dot product with the image, we are getting an activity and that activity will be following a sparse distribution. And with this simple idea, what they showed is that with just simply these two constraints, we can actually show that it is the edge detection that is done by simple cell receptive fields in the primary visual cortex as you know and orientation selectivity. It is that can easily that actually emerges from theoretical principles and we know from measurements in the primary visual cortex as you have known or we have discussed in the beginning is that they have receptive field very much like what you are seeing here that is you have different kinds of orientations, oriented bars. So you have white oriented bar with two dark bands on the side. Here is another orientation and two dark bands on the side.

Similarly the other way also with the dark band in the middle of one orientation and bright bands on the side. So and they also quantified all the spatial frequencies that they obtained from larger and larger sets of data and they found that the oriented bars that they were seeing as the receptive fields of this theoretical neurons if you will very much matched what we see in the visual cortex in the primary visual cortex with simple cells. So that is they span the entire spatial frequency range and also the entire orientation direction orientation tuning or rather orientation angle that goes from 0 to 180. So all of these were covered. So with this we can see that okay we can approach these kinds of problems from theoretical ways and with very very nominal assumptions about the system and these assumptions are indeed meaningful although there is no concrete proof for using those assumptions but it is simply what you expect should be happening.

So the thing is that so this was a remarkable result. It was a paper in Nature in 1996 and then there were follow up papers from the same group and some from others but this did not go much beyond the simple cell emergence of simple cell receptive fields. So that means that there is still much more in terms of what are the things that need to be optimized. Here we are talking of only the two things. You can now think of many other kind of factors that need optimization in terms of resources that are being used.

For example, minimum wiring, minimum sort of complexity of creating the network, whatever resources are being used to create the network, those resources also need to be minimized. So those become slightly more difficult and the formulation of the problem is also more difficult but it is not impossible and we hope that with this last week lectures you will be developing analytical or this kind of a theoretical interest. Some of you who are theoretically minded more so than others who may be more experimentally minded. Both have to go hand in hand obviously because what is the, I mean if we predict something again what is the final proof that indeed this is happening.

We have to do the experiments. But once certain ideas become a fact of observations being explained by those principles, then I think gradually we will make stronger and stronger progress in principles of neural coding through theoretical means. So with this I will stop.