

Research Interests

I am interested in how information theory and machine learning can be used to further the understanding of the functions of the brain, in particular the visual system. Information theory can help to detect the characteristics of stimuli, synthetic or natural, which are most important to the brain. That is, those characteristics which carry the majority of the mutual information. Machine learning can be used to build a model, which attempts to use these characteristics in a similar way to the brain. The accuracy of these models offer an insight into the functions of the brain.

Information Theory & Machine Learning applications

Stimulus-response model parameters

In computational neuroscience, the response of a neuron, and the stimulus which elicits that response, can be treated as random variables. This allows the use of information theory for analysis of the response and its interaction with the stimulus. This can be useful when analysing the ‘goodness’ of stimulus-response models used in decoding. For example, a stimulus-response model must make some assumptions about which response parameters are most important for carrying information. The importance of a parameter can be estimated by, firstly, assuming the parameter is relevant to the response, then measuring the mutual information between the stimulus and the response, $I(S; R)$, and then comparing this to the mutual information when the parameter is ignored. If much of the information remains, the parameter is not that important after all.

Nirenberg et al. followed this procedure in [1] when examining the encoding capabilities of a population of neurons, versus the encoding capabilities of those neurons individually. The parameter tested was the correlation of spiking across the population. It was found that 90% of the mutual information was retained when the correlation was ignored. This indicated that retinal ganglion cells act mostly independently when encoding a stimulus. The loss in 10% of the information may be accounted for by Schneidman et al. in [2], where a maximum entropy model constrained by weak pairwise correlations between neurons proved to be a very good model.

Local field potentials

Some work has also been done analysing the local field potentials (LFP) of neuronal populations. In [3], Mazzoni et al. discovered that information was carried in two frequency bands, one low and one higher. Each band carried different information about the stimulus. But the low frequency band carried information which was not carried by the spike

response. The higher frequency band carried information similar to that carried by the spike response.

An obvious follow up to this research would be to find the aspects of the natural image stimulus used in [3] that the LFPs were encoding. Then see if there is any pattern in the stimuli encoded by spike response vs the stimuli encoded by LFPs across multiple areas of the brain.

Model testing

Information theory can also be used to test models. Given a model, its parameters can be set based on the principle that biological systems should have as little redundancy as possible. This principle was presented by Barlow in 1961 [4]. The results of this model can then be compared to biological data in order to test the accuracy of the model. Even if the model is not good, the testing will offer some insight into the biological process being modelled.

Challenges in Information Theory

One issue with information theory analysis in computational neuroscience is the limited number of responses available for any given stimulus. The number of responses recorded is generally much less than the number of potential responses. Even if the response is encoded in a simplistic way, e.g. a series of spike counts, with one count for each trial, if the maximum spike count observed is 100, then there are 101 possible responses. In order to record each of these a number of times, several hundred trials would be required. Usually, executing several hundred trials is not practicable. This leads to an incorrect estimation of the conditional probabilities $P(r|s)$, where r is a certain response, and s is a certain stimulus, which has a knock on effect leading to an upward bias in mutual information. This is called the limited sampling bias problem [7]. Methods have been developed for estimating this bias, and therefore correcting it [8].

Research has also been done into the more fundamental problem of how the probabilities required for information theory analysis are defined. This work has also addressed the limited sampling bias problem. In [9], the author takes a topological approach to calculating probabilities, this helps to overcome the sampling bias. I'm interested in this kind of approach as it appeals to my background in mathematics. It is easy to believe that there maybe other ways of calculating an unbiased estimate from limited data than simply correcting the estimate calculated from 'plugged-in' data.

MSc Dissertation: How informative are retinal ganglion cells about visual stimuli?

My MSc Dissertation was in the area of computational neuroscience applied to the visual system. It was entitled How informative are retinal ganglion cells about visual stimuli? The aim was to measure the mutual information between some stimuli and the response to those stimuli of some retina preparations. It was hoped that the retinal ganglion cells (RGCs) could be classified based on their mutual information. The retina was presented with a full field flash stimulus initially, that is, alternating bright and dark light presented for equal duration. A group of cells with mutual information of 1 bit stood out from the rest. So, classification seemed possible in this case.

The second class of stimulus used was an animation of black and white bars moving perpendicular to their length. The the orientation was rotated by 45° into 8 different combinations of orientation and direction. It was expected that some highly directionally selective cells would be found based on previous findings [6], and that this selectivity would be reflected by a mutual information around 2 bits. This was not observed. It was found that some cells had enough mutual information to approximate a certain direction, but these cells were not finely tuned. There were no cells found which could distinguish all of the directions. Most cells with greater than 1 bit of mutual information appeared to be attuned to an orientation rather than a direction because these cells had a low directional selectivity index.

It was hoped that the cells which were directionally selective could be identified by their mutual information. This was not the case. Those cells which responded to one direction much more strongly than all the other directions could only distinguish one direction from all the others. This meant that their mutual information was quite low < 0.5 bits.

References

- [1] S. Nirenberg, S.M. Carcieri, A.L. Jacobs, P.E. Latham, *Retinal ganglion cells act largely as independent encoders*. Nature 411, 689-701, (2001)
- [2] Elad Schneidman, Michael J. Berry, Ronen Segev, William Bialek, *Weak pairwise correlations imply strongly correlated network states in a neural population*. Nature 440, 1007-1012, (2006)
- [3] Alberto Mazzoni, Nikos K. Logothetis, Stefano Panzeri, *The information content of Local Field Potentials: experiments and models*. arXiv:1206.0560, [q-bio.NC], (2012)
- [4] H.B. Barlow, *Possible principles underlying the transformations of sensory messages*. Sensory Communications, 217-234, (1961)
- [5] Quoc V. Le, Marc Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S. Corrado, Jeff Dean, Andrew Y. Ng, *Building high-level features using large scale unsupervised learning*. Proceedings of the 29th International Conference on Machine Learning, (2012)
- [6] Shijun Weng, Wenzhi Sun, Shigang He, *Identification of ON-OFF direction selective ganglion cells in the mouse retina*. Journal of Physiology 562, 915–923, (2005)
- [7] Stefano Panzeri, Alessandro Treves *Analytical estimates of limited sampling biases in different information measures*. Network: Computation in Neural Systems 7, 87–107, (1996)
- [8] Stefano Panzeri, Riccardo Senatore, Marcelo A. Montemurro and Rasmus S. Petersen, *Correcting for the Sampling Bias Problem in Spike Train Information Measures*. Journal of Neurophysiology 98, 1064-1072, (2007)
- [9] Houghton C. 2015, *Calculating mutual information for spike trains and other data with distances but no coordinates*. Royal Society Open Science 2, 140391, (2015)