SpikeNet : Real-time Image Processing with a Wave of Spikes.

Simon J. Thorpe, Rudy Guyonneau, Nicolas Guilbaud & Rufin VanRullen

Centre de Recherche Cerveau & Cognition UMR 5549,
133, route de Narbonne, 31062, Toulouse, France, and
SpikeNet Technology S.A.R.L., Ave de Castelnaudary, 31250, Revel, France
simon.thorpe@cerco.ups-tlse.fr

Abstract
SpikeNet is a image processing system that uses very large scale networks of asynchronously firing neurons. The latest version allows very efficient object identification in real-time using a video input, and although this specific implementation is designed to run on standard computer hardware, there are a number of clear implications for computational neuroscience. Specifically, SpikeNet demonstrates the plausibility of visual processing based on a single feed-forward pass and very sparse levels of firing. Above all, it is one of the very few models compatible with the severe temporal constraints imposed by experimental data on processing speed in the visual system.

Introduction
Originally developed as a research project for investigating the computational properties of large-scale networks of asynchronously firing neurons, SpikeNet has, in the past two or three years, emerged as a commercially viable image processing system. The source code of the original research version, developed by Arnaud Delorme during his doctoral thesis, can now be downloaded under GNU licence from his website. For the commercial version, developed by SpikeNet Technology SARL in close collaboration with Simon Thorpe's research group at the Brain and Cognition Research Centre in Toulouse, priority was given to providing a software package that was capable of reliable real-time image processing in real-world situations. Nevertheless, both versions share the same underlying computational principles, and in this presentation we will attempt to demonstrate why the encouraging results obtained with SpikeNet have major implications for Computational Neuroscience.
Both versions test the idea that high level visual processing tasks can be performed under conditions where each neuron in the system only gets to fire at most one spike, and where the percentage of neurons that actually emit spikes is kept to a strict minimum. The idea of processing with one spike per neuron dates back to the late 80s when it was argued (Thorpe & Imbert, 1989) that the speed of processing in the visual system (and in particular the fact that face selective neurons in the temporal lobe can start firing as little as 80-100 ms after stimulus onset) implies that the underlying processing can be performed with as little as 10 ms of computation time per processing step. Given that firing rates of cortical neurons rarely exceed around 100 Hz, the implication was that few neurons will get to fire more that one spike.
Obviously, in the real nervous system, it is (at least for the foreseeable future) impossible to prevent neurons from firing multiple spikes. As a result, it is unlikely that it will be possible to *prove* that the nervous system can perform high level visual tasks with only one spike per neuron. However, by building a synthetic system such as SpikeNet in which multiple spiking is prevented we can ask just

how much visual processing can be achieved under such conditions

Performance of SpikeNet and implications for visual computation

To gain some idea of the sort of performance that can be obtained with SpikeNet, we strongly recommend experimentation with the downloadable demonstration version that can be obtained from SpikeNet Technology's website at . Briefly, given an image or video sequence as an input, SpikeNet can be trained to recognize and localize specific visual forms within natural images. Hundreds or even thousands of different visual prototypes can be learnt, and processing time is directly proportional to the number of prototypes to be tested and the size of the input image. By keeping the percentage of active neurons at a minimum, accurate identification is possible even under real-time conditions, at rates that can exceed 200 Megapixels per second on a Pentium 4 based machine at 2 GHz. Obviously, the processing rates that can be achieved on a computer are difficult to compare with processing in real neural networks. Nevertheless, there are a number of more general conclusions that can be drawn.

Single spike coding is viable. The first point is that sophisticated visual processing with just one spike per neuron is clearly possible, despite that fact that with only one spike, traditional coding schemes based on determining the firing rates of individual neurons are ruled out.

Pure feed-forward mechanisms are computationally powerful. Although SpikeNet can include both horizontal and feedback connection patterns, the current version does not use them. Despite this, accurate identification is possible even with noisy images and at low contrasts**.** Clearly, the initial feed-forward wave of processing is capable of considerably more than is conventionally assumed.

Sparse coding is very efficient. One of the main reasons for the speed of SpikeNet lies in its very sparse coding scheme. Typically, we have found that only 1-2% of neurons in any given processing stage need to fire in order to allow identification. The key is to use a coding scheme in which the most strongly activated neurons fire first (Rank Order Coding) since this guarantees that decisions are made as quickly as possible.

Image segmentation is not required for high level identification. One of the most striking features of SpikeNet is that there is nothing even remotely like image segmentation going on. Everything is done by using large numbers of neurons tuned to diagnostic combinations of features that will fire as soon as there is enough evidence to allow activation. It could be that the traditional view that the first step in processing requires scene segmentation is a major error, and that intelligent segmentation involves feedback that occurs only one the initial feed-forward pass has been completed.

Future developments

The processing architectures used by SpikeNet are still a long way from those used by biological vision and future work will be aimed at reducing the gap. For example, SpikeNet does not have the equivalent of separate ventral and dorsal pathways specialised for object identification and localisation. Instead, there is a retino-topically organised map of neurons for each object or feature constellation that needs to be identified. For applications, this is actually quite useful, because the system automatically provides about the *xy* coordinates of each identified object (unlike object selective neurons in inferotemporal cortex that have only limited spatial selectivity). However, there is a very high cost in terms of the number of neurons required. Future versions will try are use a more biologically realistic strategy that almost certainly will allow a major reduction in the number of neurons required. Nevertheless, the biological reverse engineering approach used in SpikeNet has already proved remarkably successful and a number of important computational issues have already been addressed using this sort of approach.