

# ULTRA-RAPID VISUAL CATEGORISATION USING ONE SPIKE PER NEURON

Simon J. Thorpe

Centre de Recherche Cerveau & Cognition  
133, route de Narbonne, 31062, Toulouse, France  
[thorpe@cerco.ups-tlse.fr](mailto:thorpe@cerco.ups-tlse.fr)  
[www.cerco.ups-tlse.fr](http://www.cerco.ups-tlse.fr) [www.spikenet-technology.com](http://www.spikenet-technology.com)

## 1. Introduction

In 1989, I argued that the response latency of face-selective visual responses in high order visual areas such as the primate inferotemporal cortex poses severe problems for almost all models of visual processing that rely on iterative processing (Thorpe & Imbert, 1989). Such neurones start firing 80-100 ms after stimulus onset, which leaves only about 10 ms for processing at each stage of the visual pathway. The original argument could be criticised on the grounds that face processing could be a special case and that in such neurophysiological studies, the stimuli are often highly predictable. However, more recent experiments have shown that very fast processing also occurs in much more challenging visual tasks where subjects must decide whether a briefly flashed natural image contains a target category such as an animal or a means of transport (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001). The most serious implication of such experiments is that very sophisticated visual processing must be possible under conditions where individual neurones only have time to emit one spike. Clearly, this rules out most coding schemes that require the firing rates of individual neurones to be determined with precision. However, it leaves open a number of options in which the precise timing of spikes across a population of cells is used.

## 2. Rank Order Coding

In Rank Order Coding, we make use of the simple fact that more strongly activated cells will tend to fire first. As a result, one can learn a lot about the stimulus simply by looking at the order in which cells fire (Thorpe & Gautrais, 1998). Indeed, Rufin VanRullen has recently shown that this sort of coding scheme provides a very efficient way to transmit information from the retina to the cortex - far more efficient than conventional rate coding based schemes. Specifically, he found that detection of animals in natural images is usually possible when only 1% or so of retinal ganglion cells have fired one spike (Van Rullen & Thorpe, 2001). Although little direct experimental evidence is available on this point, it seems likely that this would occur within a short enough time window to be compatible with ultra-rapid visual categorisation.

Other work has shown that simple cortical circuits involving feed-forward shunting inhibition can be used to make neurons at later stages selective to the order in which their inputs fire. This can be used to produce orientation selectivity in neurons in striate cortex, but more interestingly, it can be used to obtain neurons capable of detecting faces in natural scenes (Van Rullen, Gautrais, Delorme, & Thorpe, 1998). Recently, it has even been demonstrated that a simple feed-forward architecture that makes use of asynchronously firing neurones and rank-order sensitivity can be used to identify and localise faces in natural images (Delorme & Thorpe, 2001). One of the remarkable features of such models is that

they are naturally very resistant to changes in image contrast. This follows naturally from the fact that while firing rates and actual onset latencies will vary a lot when the contrast of the input image changes, the order in which cells fire tends to be very robust. Contrast invariance is a feature that appears to occur naturally in biological vision systems but which causes major problems in typical machine vision systems that often rely on the grey-scale values of pixels in the input image.

### **3. Implementation in SpikeNet**

Our initial experiments with networks of asynchronously firing spiking neurons were made possible using SpikeNet, a program developed by students in our laboratory. The first version was developed by Jacques Gautrais for his thesis in 1997 and this was followed by a second version written by Arnaud Delorme (Delorme, Gautrais, van Rullen, & Thorpe, 1999). Given the remarkably good performance obtained using SpikeNet, it was decided to develop a commercial version for use in image processing applications. This latest version has been developed by a start-up company, SpikeNet Technology S.A.R.L. ([www.spikenet-technology.com](http://www.spikenet-technology.com)) with the aid of investment from regional development agencies. Initial results are very promising because it is clear that this sort of spike-based neural network simulation system can be implemented very efficiently using conventional computer hardware. For example, the current version is capable of identifying several hundred different face images per second on a standard Pentium 4 based computer running Windows. Performance is sufficiently good that we expect to be able to do real-time image analysis using full-size video inputs at up to 25 frames per second within the next few months. Furthermore, we are able to simultaneously test for thousands of different visual forms with only small requirements in terms of memory.

### **4. Implications for coding**

Although this is not the appropriate place for a commercial presentation, it is nevertheless interesting to comment on the reasons behind the remarkably good performance obtained with such systems. Essentially, the "trick" is to use only a very limited number of spikes to do the processing. Firstly, the way the simulator works means that no neurone can fire more than one spike - effectively ruling out conventional rate based coding schemes. However, by propagating spikes in the order of their activation levels we are guaranteed that the most useful and salient information is sent first. The use of such a strategy means that we can progressively reduce the proportion of cells that fire until we are left with just the strict minimum required for recognition. Both the results of Rufin VanRullen's model and our SpikeNet simulations imply that only a very small proportion of cells need to fire - maybe only 1%. Clearly, when such sparse codes are implemented on a computer, there is a large reduction in the amount of computation required and this is one of the major factors behind SpikeNet's efficiency. But similar arguments apply in the case of biological hardware - generating action potentials costs energy and so natural selection will favour coding schemes that minimise the number of spikes needed. Asynchronous spike propagation proposed here will have just this feature since the whole aim of the system can be described in terms of trying to get as much computation done with the strict minimum number of spikes.

One final comment concerns the fact that the current versions of SpikeNet only use a small subset of the various coding options offered by networks of spiking. In recent years, more and more researchers have become interested in the computational implications of spike-based computing (Maass & Bishop, 1999; 1996; Thorpe, Delorme, & VanRullen, 2001). It seems likely that in the future, the importance of spikes for both biological and artificial neural networks will become more and more significant.

## References

- Delorme, A, Gautrais, J, van Rullen, R & Thorpe, S (1999). "SpikeNET: A simulator for modeling large networks of integrate and fire neurons" *Neurocomputing*, 26-7, 989-996.
- Delorme, A & Thorpe, S J (2001). "Face identification using one spike per neuron: Resistance to image degradations" *Neural Networks*, in press.
- Fabre-Thorpe, M, Delorme, A, Marlot, C & Thorpe, S (2001). "A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes" *J Cogn Neurosci*, 13, 171-180.
- Maass, W & Bishop, C M. (1999). *Pulsed neural networks*. Cambridge, Mass.: MIT Press.
- Rieke, F, Warland, D, Ruyter van Steveninck, R & Bialek, W. (1996). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- Thorpe, S, Fize, D & Marlot, C (1996). "Speed of processing in the human visual system" *Nature*, 381, 520-522.
- Thorpe, S J, Delorme, A & VanRullen, R (2001). "Spike-based strategies for rapid processing" *Neural Networks*, (in press).
- Thorpe, S J & Gautrais, J. (1998). Rank Order Coding. In J. Bower (Ed.), *Computational Neuroscience: Trends in Research 1998* (pp. 113-118). New York: Plenum Press.
- Thorpe, S J & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer & Z. Schreter & F. Fogelman-Soulié & L. Steels (Eds.), *Connectionism in Perspective*. (pp. 63-92). Amsterdam: Elsevier.
- Van Rullen, R, Gautrais, J, Delorme, A & Thorpe, S (1998). "Face processing using one spike per neurone" *Biosystems*, 48, 229-239.
- Van Rullen, R & Thorpe, S J (2001). "Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex" *Neural Computation*, 13, 1255-1283.
- VanRullen, R & Thorpe, S J (2001). "Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artefactual objects" *Perception*, 30, 655-668.