- Friday, September 18, 1998
  The Correspondence Problem

1.0 Before we get to the main body of this report, I want to mention that the recent NSERC site visit seemed to go off quite well. In the course of the visit, the visitors mentioned that they had spent some considerable time reading through everybody's web pages to prepare themselves for the visit. Hope they liked yours!
2.0 I hope everyone is making sterling efforts to get suitable people to apply for our two job openings (see web site), especially the senior, ORDCF-funded position. More on who has applied etc... when we have a clearer picture-- perhaps we should have a meeting soon to discuss progress. Guess where potential applicants will be going to find out more about YORKVIS...
3.0 Next YORKVIS talk is by Doug Munoz from Queens. He will be giving TWO talks on Friday 16th October: one at 10am to YORKVIS (room 061 BSB "Title not announced but I am guessing something on the Superior Colliculus") and another at 2pm as the YORK PSYCHOLOGY DEPARTMENT COLLOQUIUM (room 291 BSB "Using eye movements to probe brain function and dysfunction")
4.0 The notes of Harry Sperling's talk (date; web site) have been updated slightly in the light of his comments on the previous version. All the other reports have now been approved by the relevant speakers.
5.0 The talk on Sept 18th was given by David Fleet on modelling the neural basis for solving the correspondence problem. Sorry I have been so tardy in writing it up. The correspondence problem is, when combining the images coming from the two eyes into a single perception, how do you know which part of the left eye's image goes with which part of the right eye's image?
5.1 Fleet took Ohzawa's model (Ohzawa et al, "Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors." Science. 1990 249:1037-41). as his starting point. In Ohzawa's model Hubel and Weisel's simple cells are redescribed as "binocular linear neurones" and their complex cells become "binocular energy neurones". Binocular linear neurones are modelled as the sum of two receptive fields, one in each eye, each with a set of delineated excitatory and inhibitory regions. Binocular energy neurones (complex cells) are modelled as the sum of rectified linear neurones. 5.2 The response of these binocular energy cells is determined by the sum of the energy falling on the component receptive fields in the two eyes plus an interaction term which depends on the difference in phase between the left and right the monocular responses (the response phase depends on the exact position of the stimulus in the receptive fields). Freeman and Ohzawa recorded from binocular energy cells and varied the phase of a grating presented in one eye relative to the phase of a second grating presented simultaneously to the other eye. They found a preferred phase for each cell. BUT they also found that varying the contrast of the stimulus in one eye surprisingly produced almost no change in the response of the cell. The contrast is the driving energy of the stimulus, so any model based on energy is challenged by this latter finding.
5.3 To rise to this challenge, Fleet took the basic model described by Ohzawa but added to the left and right eye's input a "gain control" box that could boost the response to compensate for a low contrast input. This is not quite as arbitrary as it sounds as there are several perceptual reasons for postulating gain control in the visual system. Gain control was also added to the output of the binocular energy neurones.
5.4 This modified model was then tested against published data of the responses of simple and complex cells (eg. Anzai, 1995 "Contrast coding by cells in the cat's striate cortex: monocular vs. binocular detection." Vis Neurosci. 1995 12:77-93.) and their variations with contrast. The model fitted the data nicely.
6.0 Fleet now considered models for disparity detection, that is how are the small differences between the two eyes (disparities) that provide important information for depth perception coded?
6.0.1 One model, originating with Hubel and Weisel, is that the receptive fields are in slightly different locations in the two retinae so that for a feature to fall squarely on both fields at the same time, it has to have the appropriate disparity. Hence the cell will only fire maximally in the presence of a particular disparity.
6.0.2 A second model, proposed by De Angelis, Ohzawa and Freeman (Perception 1997) suggested that the boundaries of the receptive fields might stay the same but that the inhibitory and excitatory regions within the receptive field might be shifted.
6.0.3 Fleet proposes a hybrid model with both the above features.
6.0.4 But how to distinguish between these three models?
6.1 Fleet implemented his hybrid disparity-detection model into his gain-control-modified, binocular-energy model. There are two variables in the model corresponding to the shift in fields (H&W model) and the shifts in regions within the field (D, O & F model). By seeing what these values need to be to fit experimental data, the model can indicate the relative importance of these two factors.
6.1.1 When tested against owl visual wulst data (from Wagner and Frost; Comp Physiol. 1994) the model suggested a position shift of 1.8 degs as well as a phase shift of -0.5 rads thus supporting Fleet's hybrid model. I am not sure why the model was tested on birds and not monkeys...
6.2 So the model is looking quite comprehensive: but can it find correspondences? The answer is no. The model cells are simply adding energy. If there is an ambiguity- more than one possible match - the model cannot choose one over the other. That is these binocular energy cells being modelled here are not disparity detectors.
6.3 But Fleet described a way that cells could be added together to create pools that could do the job. Pooling could occur over local spatial neighbourhoods, over different orientation tuned cells and cells tuned to difference scales. He showed some simulations of the way such pools could be used to identify certain values of disparity in a stereophotograph.
7.0 The talk on 28 Sept 98 was by Gabriel Robles De La Torre and was entitled "Internal models of virtual objects"
7.1 De La Torre's premise was that to make skilled manipulations of an object, you need information that is not present in the stimulus. That is you need outside knowledge. The example he described at length was pushing or pulling a piston that was immersed in some viscous oil. But

this point is true for any manipulation such as opening a door or operating a gear lever on a car.
7.2 To apply this knowledge usefully, he proposed, requires an internal model or representation of the system under consideration.
7.3 Since most systems can be modelled by the differential equations of systems dynamics, this could be taken to imply that the equations, or their equivalent, are internalized.
7.4 This principle applies to virtual objects (eg an interactive animation on a computer screen) in exactly the same way. The only difference would be the sort of feedback one obtained, or perhaps just the medium of the feedback, whilst learning the manipulation or task.
8.0 De La Torre did some experiments in which he trained people to push buttons to manipulate a computer animation that simulated a piston in oil. By looking at their improvement in performance he attempted to identify the strategy that they were using (or internalizing).
8.1 In common with the strategy that seems to be used for many physiological tasks, subjects seemed to adopt an energy-minimizing strategy.
8.2 Was this strategy achieved by learning the differential equations or by some kind of temporal calibration? To investigate this, having trained subjects on the single virtual piston, De La Torre cleverly changed the simulation into a double piston. The minimal errors subjects made were more compatible with having internalized the equations to describe the system (which were still valid) rather than calibrating time, which would have required a whole new calibration.
8.3 More detailed discussion on the role of feedback loops versus internalized equations was unfortunately cut short by time pressure..

David Fleet