A combined pyschophysical–modelling investigation of the mechanisms of tactile picture perception

Andrew P. Davison*§, Pierre Yger§, Jason S. Chan†, Fiona N. Newell† and Yves Frégnac§

§Unité de Neurosciences Intégratives et Computationnelles, Centre National de la Recherche Scientifique, Gif sur Yvette, France

†Department of Psychology, University of Dublin, Trinity College, Dublin, Ireland

*Corresponding author: davison@iaf.cnrs-gif.fr

Introduction

Cross-modal interactions between sensory modalities are of vital importance in normal perception, particularly when the characteristics of the object being perceived are the same or similar in both modalities. For example, both visual and tactile object recognition are based on the spatial arrangement of certain basic features such as edges and textures. There is an accumulating body of evidence that tactile/haptic perception of complex (two- and three-dimensional) objects involves visual imagery [6, 8, 9]. Brain regions in which visual and haptic information may be integrated have been identified [1, 2, 4, 7], but the mechanisms by which haptic and visual-mental-image information is integrated are unknown.

To address this deficit, we have pursued a joint experimental-modelling approach. We have developed an experimental protocol to test the performance of human subjects in a two-choice forced decision task which requires haptic perception of a scene composed of two objects in order to determine their spatial relationship. The objects are differentiated by surface texture, and the difficulty of the task may be varied by controlling the difference in texture between the objects. The scan-paths taken by the subjects during the task are recorded. We have further developed a computational model, based on the principles of visual imagery during tactile processing outlined above, of the computational steps that we hypothesise are performed by the human subjects in performing the task. The model receives the same tactile input and follows the same scan paths as the human subjects. We are able to make a detailed quantitative comparison between model simulation and human experimental results in order to elucidate the role of different hypothesised exploration strategies and perception mechanisms. We present here the experimental results and preliminary modelling results.

Experimental methods

Subjects are presented with a series of pictures made up of raised dots, in a horizontal plane. Each picture consists of two rectangles, one of which partially-obscures the other (Fig. 1). The task is to scan the picture surface with one finger and determine which of the rectangles (‘upper’/‘far’ or ‘lower’/‘near’) is on top of the other.

Pictures of size 60×40 tactile pixels (taxels) are presented on a virtual tactile display (VTD)[5]. The subject’s finger rests on a moveable carriage, and can feel a 3×3 taxel subset at any one time. As the carriage is moved around the picture, the dot states (raised/lowered) are recalculated each time the finger crosses a taxel boundary, based on the mean dot density (probability of a dot being raised). The VTD records the scan path taken by the finger, and the state of the taxel array. The subject must respond within 60 s.

To generate a psychometric curve, the difference in dot density between the two rectangles is varied. The density of dots in the upper rectangle, d_{U}, is varied from $d_{\min}$ to $d_{\max}$, and that in the lower rectangle, d_{L}, set to $d_{\max} + d_{\min} - d_{\mathrm{U}}$. We define $\lambda = \pm |d_{\mathrm{U}} - d_{\mathrm{L}}| / (d_{\max} - d_{\min})$. λ is negative if the lower rectangle is on top, and positive if the upper rectangle is on top, irrespective of which rectangle is the denser. In each of four experiment blocks, 39 stimuli are presented, six for each non-zero value of λ, and three for $\lambda = 0$. We used $d_{\max} = 0.8$ and $d_{\min} = 0.2$ dots/taxel. The frequency with which the participant responds ‘upper’ is plotted as a function of λ.

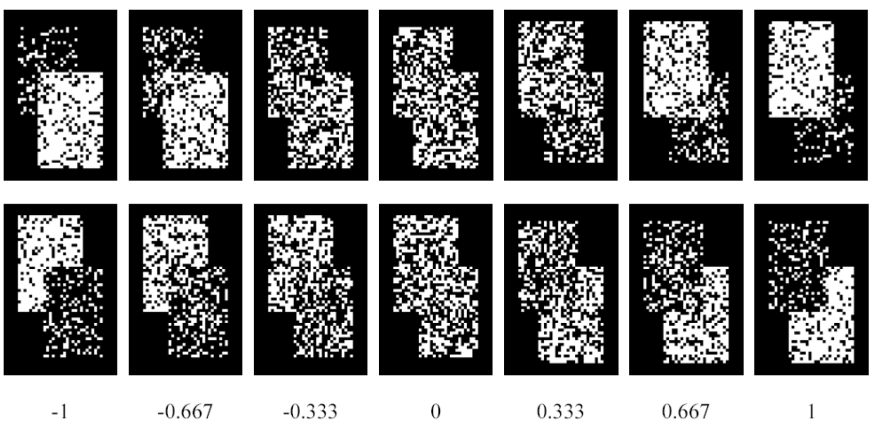
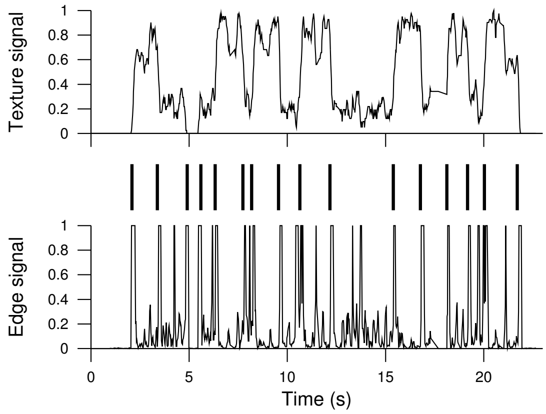


Figure 1: Snapshots of haptic stimuli. White points represent raised taxels, black dots represent lowered taxels. Numbers show the value of λ. The exact pattern of taxels changes each time the carriage is moved; only the probability of a taxel being raised in each rectangle is constant. For each value of λ there are two possible stimulus conformations.

Figure 2: Example tactile signals. The taxel states and finger location re recorded during the human psychophysical experiments. These data are used to construct input signals for the model. The texture signal is derived from a spatial and temporal average of the taxel state. The edge signal is derived from the texture signal. The solid vertical bars between the two plots show the actual edge locations. The edge signal reproduces these quite faithfully, although with some additional spurious edges.



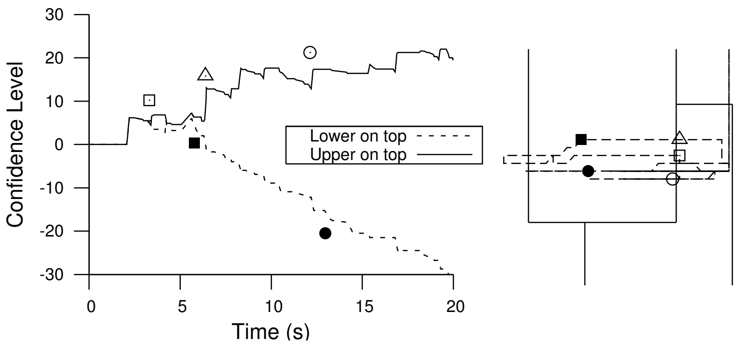


Figure 3: Evolution of the confidence level for each hypothesis during the scan. The scan path is shown at right. Open symbols mark points where the tactile signal was in good agreement with the 'upper rectangle on top' hypothesis, giving a sharp increase in the confidence level for that hypothesis, and a corresponding decrease for the other hypothesis, as the tactile signal does not match the expectation in that case. Filled symbols mark points where an edge is expected for the 'bottom rectangle on top' hypothesis but none is found. The confidence level for this hypothesis therefore declines while the other confidence level is unchanged. Note that to clearly show the functioning of the model, most sources of noise were removed for the purposes of this figure ($\sigma_V = 0$, $\sigma_{xy} = 0$).

Model description

The subject has prior knowledge about the possible rectangle configurations, and so only two hypotheses are necessary: 'upper rectangle is on top' and 'lower rectangle is on top'. The subject is assumed to alternate rapidly between the two hypotheses, generating each time a corresponding mental image. It is assumed that at some stage of the ventral visual pathway, the activities of neurons that are responsive to particular features found in the expected shape (horizontal edges, vertical edges, corners) are proportional to the expectation that the preferred features will be found within their receptive fields. An 'attentional spotlight' controlled by the current position of the finger within the mental image space gates the tactile input to this population, such that only those neurons whose receptive fields contain the current finger position can give an output. An additional 'spotlight' based on the current direction of motion ensures that only those neurons responsive to horizontal or vertical edges, as appropriate, are activated.

Simulated tactile signals (Fig. 2) are derived from spatial and temporal averaging of the taxel states, with edges between regions of different mean density detected as sudden changes in this averaged signal.

Each time the path moves into a different taxel, the expectation E is compared to the tactile signal T. The agreement, A, between tactile signal and expectation is calculated as $A = T \times E$ if $T \geq \theta_T^A$ and is zero otherwise, where θ_T^A is the threshold above which tactile changes are considered significant. The disagreement, D, is calculated as

$$D = \begin{cases} E & \text{if } T < \theta_T^D \text{ and } E \geq \theta_E^D \\ T & \text{if } T \geq \theta_T^D \text{ and } E < \theta_E^D \end{cases}$$

and is zero otherwise. In other words, if a feature is expected but none is present, the disagreement is given by the size of the expectation, and if a feature is found where none was expected, the disagreement is given by the strength of the tactile signal. The confidence level C is then changed according to $C = C + kA - D$ where k controls the balance between agreement and disagreement in determining the confidence level. $k = 10$ for the simulations shown here, chosen so that in general by the end of the scan one of the confidence levels is positive and one negative.

Most of the human subjects restrict their scan paths to the most informative region of the figure, suggesting that they have a good memory for the location of the rectangles within the scene. However, this memory is unlikely to be perfect, so to generate the mental images we take the actual coordinates of the rectangle edges and jitter them by adding a random offset taken from a normal distribution with mean 0 and standard deviation σ_{xy} (4 taxels in the simulations shown here). Furthermore, it has been shown that for humans, in the absence of visual feedback, knowledge of hand position tends to drift over time [3]. Therefore, noise (Gaussian, mean zero, standard deviation $0.2V$) is added to the model's representation of its current velocity, V.

The evolution of the confidence levels during a typical scan path is illustrated in Figure 3.

Results

Subjects invariably adopted a stereotyped scanning pattern, consisting of either mainly vertical or mainly horizontal movements (Fig. 4). Four of the subjects mostly restricted their scanning to the region in which the rectangles overlapped, since this is the

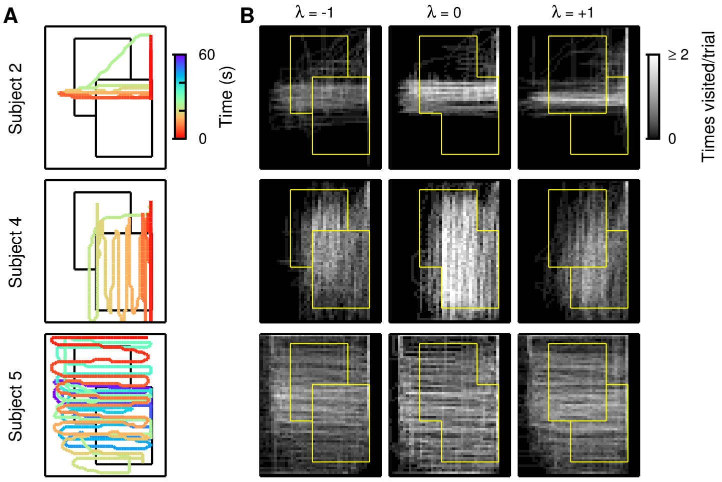


Figure 4: Subjects adopt stereotyped scanning strategies. (A) Typical scan paths for three subjects. Subjects 1 and 3 are not shown, but had a similar strategy to Subject 2. The colour-scale represents the time elapsed during the trial. (B) Average scan paths for $\lambda = +1, 0, -1$, averaged over all trials at a given λ. The grey-scale value represents the mean number of times a given taxel was visited during one trial.

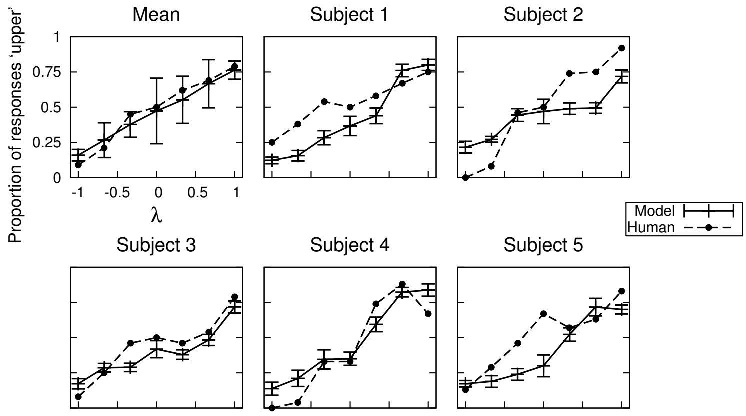


Figure 5: Comparison of model and human psychometric curves. For the mean across subjects, the model was run once. The model curve shows mean and standard deviation across subjects. Individual subject model curves show mean and standard error ($n = 10$ runs, different random seeds). The mean psychometric curves were approximately matched by setting the uncertainty in the mental image (σ_{xy}) to 4 taxels. The curves for individual subjects match quite well for Subjects 3 and 4, but not for the other subjects.

most informative region for performing the task.

The statistics of the scan paths show a clear dependence on λ, with mean scanning speed being lower and response time for λ nearer to zero (smaller difference in dot density and hence greater difficultly in the task).

Graphs of the proportion of responses that are "upper rectangle" as a function of λ (psychometric curves; Fig. 5), demonstrate that the subjects can perform the task , and that λ has a strong influence on performance. However there is considerable variability between subjects and inconsistency within individual subjects' performances.

As a measure of how well the model responses agree with the human responses (and therefore of how well the model represents the underlying mechanisms used by humans to solve the task), we approximately match the mean-across-subjects psychometric curve for the model to that for the human subjects, and then examine the similarities for individual- subject psychometric curves. We find that for subjects 3 and 4 the agreement is quite close, but that there are large discrepancies for the other subjects. At this preliminary stage, we cannot therefore conclude that the model gives a good match to human performance.

Discussion

By developing an experimental design that provides detailed quantitative data about subject strategies and performance, we have the possibility of making a detailed comparison of model and human behaviour in order to elucidate the mechanisms underlying haptic spatial perception. Our preliminary modelling studies show that we can reproduce human behaviour when averaging over variations between individual subjects, but we cannot reliably reproduce the individual variations. In future work we intend to refine the model, for example by the addition of adaptive refinement of the mental images, in order to improve the agreement between human and model responses, and we will examine the effects of different rules in calculating agreement and disagreement of sensory input and top-down expectation. This research was supported by the European Union under the Life-Like Perception Programme, project reference IST-2001-34712 (SenseMaker)

References

[1] Amedi, A., Jacobson, G., Hendler, T., Malach, R. and Zohary, E. (2002). *Cereb. Cortex* **12**:1202–1212.
[2] Amedi, A., Malach, R., Hendler, T., Peled, S. and Zohary, E. (2001). *Nature Neurosci.* **4**:324–330.
[3] Brown, L. E., Rosenbaum, D. A. and Sainburg, R. L. (2003). *J. Neurophysiol.* **90**:3105–3118.
[4] Hadjikhani, N. and Roland, P. (1998). *J. Neurosci.* **18**:1072–1084.
[5] Maucher, T., Meier, K. and Schemmel, J. (2001). In: *Proceedings of the Sixth International Symposium on Signal Processing and its Applications*.
[6] Prather, S. C. and Sathian, K. (2002). *Brain Res. Cogn. Brain Res.* **14**:91–98.
[7] Prather, S. C., Votaw, J. R. and Sathian, K. (2004). *Neuropsychologia* **42**:1079–1087.
[8] Sathian, K. and Zangaladze, A. (2002). *Behav. Brain Res.* **135**:127–132.
[9] Zangaladze, A., Epstein, C., Grafton, S. and Sathian, K. (1999). *Nature* **401**:587–590.