Anatomy is Symmetry's Best Friend: Reflections on Modeling Baylis and Driver

John Hicks (John.Hicks@ed.ac.uk)
Division of Informatics; 2 Buccleuch Place
Edinburgh, Scotland EH8 9LW

Jon Oberlander (J.Oberlander@ed.ac.uk)
Division of Informatics; 2 Buccleuch Place
Edinburgh, Scotland EH8 9LW

Abstract

An aptitude for the detection of bilateral symmetry is a fairly prominent aspect of the human visual system. Knowledge of the reasons behind this facility is not so well established, however. Some of the behavioral data indicates that processing of symmetric and non-symmetric stimuli is undertaken in two wholly different manners (i.e. *serial* versus *parallel*). However, the interpretation of this as being due to high level cognitive preferences does not exhaust the list of possible explanations. Using a split-neural network model, we show that instead of cognitive preferences, gross morphological factors may play a large role in underwriting the ability to detect symmetry as a special case of shape perception. The earlier model is consistent with behavioral data, but Occam's razor suggests that we might prefer the newer morphological explanation.

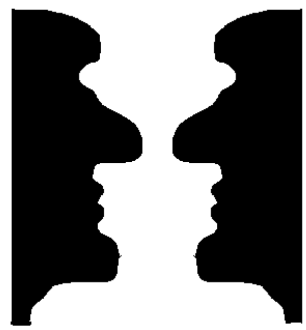

Figure 1: A familiar optical illusion whose interpretation may depend on the part-decomposition facilitated by symmetry.

Introduction

Bi-lateral symmetry is ubiquitous in nature. Such symmetry is related to biological morphology, fitness, and behavior throughout the animal kingdom (Dakin and Herbert, 1998). Thus it is not surprising that it has also been shown to be a highly salient property of the human visual system, implicated in many phenomena.

Symmetry is both a morphological characteristic and a perceptual benchmark. From recognition of a suitable mate to apprehension of a possible predator, symmetry plays an important role, being a "non-accidental" property. That is, it is unlikely that symmetry inheres in an image by chance, or when the actual image source is asymmetric. And although increasingly there is the view that symmetry detection is not only universal, but also fundamental, emerging from very low level simple processes (Dakin and Herbert, 1998; Sally & Gurnsey, 2001), there is not yet consensus about the mechanisms that underlie the facility. On the one hand, it seems that symmetry detection is a bottom-up effect of low-level filtering in early stages of the visual process. On the other, it appears to be a top-down preference for image distillation based on its exploitability for segmentation and part decomposition (Baylis & Driver, 1994; Latecki & Lakämper, 1999). Its utility in segmentation applications has caused some to comment that "the link between symmetry and segmentation curiously seems to be more than a coincidence" (ven Tonder & Ejima, 2000).

Indeed, the benefit of symmetry detection for segmentation helps promote the view that it is a worthwhile thing to be good at, for segmentation is linked to part-decomposition, which in turn could be key to figure ground separation, even aiding, for example, the interchange foreground and background in a very common visual illusion (Figure 1).

This paper deals with the specific area of contour symmetry, and its effects on human visual processing, by looking at a computational model of a specific behavioral study by Baylis and Driver (1994). The field of behavioral studies on symmetry is large; it often concerns not only contour symmetry, but internal symmetry (Hicks & Monaghan, 2001) and the effects of various filtering processes. For the purposes of this paper, however, we focus on providing a computational explanation for the differences which arise in processing symmetric and repeated shapes, as seen in Figure 2.

Behavioral Studies

Part-decomposition offers a motivation for the "symmetry is special" theory, but alone it says little about the mechanisms involved. Baylis and Driver

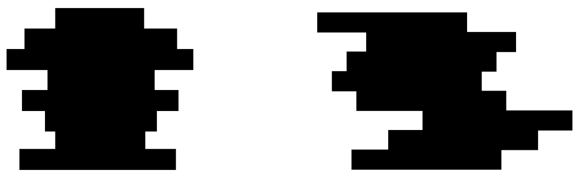

Figure 2: Stimuli: symmetric (left) and repeated (right) contours, both showing 8 discontinuities (steps) along the sides.

performed two experiments linking the perception of different shape types to distinctions in cognitive processes. In particular, the experiments aimed to elucidate the relationship between the perception of symmetry and the class of cognitive processes that are termed "parallel." In this case, "parallel" would mean that in the detection of symmetry in a two-dimensional figure, the subject *does not* engage in anything akin to a serial point-by-point comparison along the shape's contour. Baylis and Driver used a selection of perfectly symmetric shapes intermingled with shapes whose contours contained "errors" which meant a deviation from the truly symmetric form along 25% of the contour. Subjects made symmetry judgements while the experimenters varied the number of steps along the side of the shape, between 4, 8 and 16. The experimenters wanted to know whether the reaction time and error rate were significantly dependent on this variation.

It was found that symmetry was generally more quickly identified than asymmetry. This indicated directly that subjects were not involved in point-by-point search, which would always terminate earlier with erroneous examples of symmetry. Furthermore, effects of step number on subject performance were slight, and remained well within the accepted limits that define a process to be parallel.[1] Thus, the hypothesis that detection of symmetry is governed by a process impervious to increases in complexity brought about by a greater number of steps seems supported.

But what if the effects of the symmetric shapes were merely an effect of their potentially constrained nature? It might not be symmetry specifically, but redundancy in general that accounts for this semblance of parallel processing. By conducting an analogous experiment, using repetitive shapes (Fig. 2, right side), it should be possible to confirm or dismiss this confound. After all, repetitive shapes are as redundant as their symmetric counterparts, while exhibiting that redundancy via translation instead of reflection.

[1] The exception to this was when the shapes were oriented horizontally, where there was a slight effect of step number for symmetric shapes. We touch on this briefly in the discussion of our own model.

This second experiment found a significant effect for number of steps, consistent with the hypothesis that whatever process is used to judge repetition, it is effected by step count, as though it *were* a serial process. This suggests that the main difference between the two types of shapes is that in the processing of symmetry the number of discontinuities along the contour is not a significant factor, while for repeated shapes it certainly is.

Given that repetition and symmetry are equally redundant, it is clear that there must be a qualitative difference between them. The step number effect indicates a point-by-point comparison—a serial search— in the detection of repetition, which is absent from symmetry detection. But a new question arises: what is it that promotes this fast-track route for the detection of symmetry? Beneath the 'higher-level' concepts of parallel and serial processing, is there a more fundamental explanation for the fact that symmetry appears to render insignificant the relative complexity of a shape?

Modeling

This paper aims to show that this may be the case and, furthermore, that this could just as equally be the result of gross morphological aspects of anatomy as of high-level cognitive preferences. More precisely, the original assessment of the behavioral data says little about the hypothesis that what is parallel about symmetry perception is actually the 'ready-state' of the human processor to accommodate vertical symmetry.

To show this we employ a split neural network, which has previously been used as a rough correlate of the split in human visual processing, modeling reading (Shillcock & Monaghan, 2001) as well as the apprehension of the effects of symmetry in word-based stimuli (Hicks & Monaghan, 2001).

Variations in the typical architecture of neural nets often involve adding one or a number of hidden layers, vertically (Elman, 1993), or the insertion of recurrent connections. Our model employs a specific manipulation of network architecture that is not so common. Instead of devoting the entire hidden layer to the whole task upon which the network is being trained, the hidden layer can be split laterally, with each resulting half being privy to only half of the input (Shillcock et al, 2001). This split affects network performance, as shown in other modeling work. Here we apply it to the detection of symmetry in pseudo-random-block shapes.

Materials and Methods

A series of simulated neural networks, employing a back-propagation learning algorithm, were trained using sets of two-dimensional pseudo-random block shapes represented by patterns of activation. The shapes were presented to the networks through a

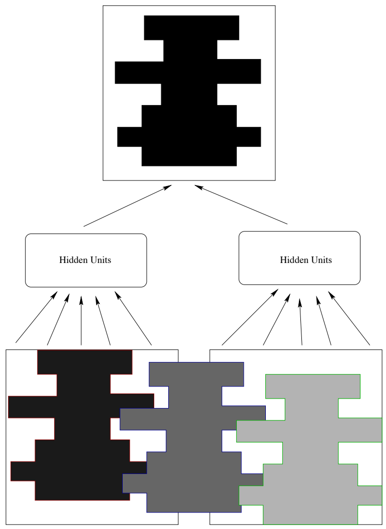

Figure 3: The split architecture network reproduces the form presented at the input, which may appear anywhere across the two visual hemi-fields.

shift invariant identity mapping (SIIM) task, maintaining the predetermined 2D block-shape of the stimuli, while moving it sequentially along the input window (Figure 3). Input nodes that fall outside the location of the block have activation zero. The vertical split in the input reflects that of the fovea and thus, as a block is repeatedly presented to the network from all possible positions across the input, it crosses from one "visual hemifield" to the other, activation being redirected to the associated hidden layer accordingly. The network is trained to recognize (represent) the shape it is being trained on.

Each stimulus set contained 60 pseudo-random block shapes, of one shape-type either all symmetric or all repetitive. For each shape type, there were three stimulus sets, with shapes having 4, 8 or 16 discontinuities along the contour (shapes with 8 discontinuities are shown in figure 2). A third class of stimulus, consisting of mixed sets, where the number of discontinuities was homogeneous, but both symmetric and repetitive shapes were represented equally, was also used in training. Due to the presentation of each pattern in all visual input positions, each stimulus accounts for 17 events in the total training set, for a total of 1020 presentation-recognition events per epoch.

After training to a predetermined number of network epochs, each net was tested with novel stimuli. The test set we focus on for the purposes of this paper contained novel blocks that were neither symmetric, nor repetitive. We were interested in seeing how the networks tended towards reproducing the type of shape they were trained on when being presented with these "random" stimuli. The metric used for gauging network performance on these test

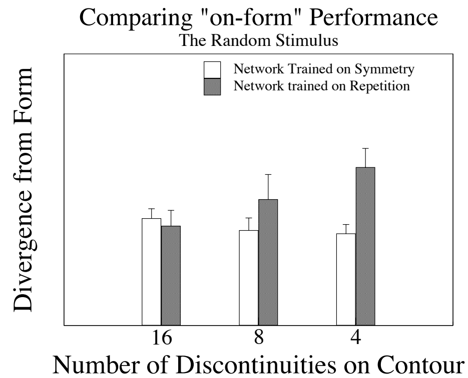

Figure 4: The compared "form" produced by trained networks, as a function of step number, under the *random* stimulus.

sets is discussed in the next section.

For all simulations the PDP++ Neural Nets software from CMU was used, running on an Ultra 5 workstation.

Results

In examining network performance on the two dimensional stimuli, we want a way of gauging the degree to which the activation at the network's output tends generally toward a given type of shape. Fortunately, the stimuli were strictly formal, in the sense that they can be defined in terms of a simple additive exemplar based function. A measure for symmetry measure is based on cancellation around the proposed axis of symmetry, while that for repetition measure but requires that activation add to a constant along arrays orthogonal to the axis of repetition (the bars that constitute the repetitive patterns are of constant length). We adopt a method of summing activations at the output so that the closer we get to the ideals, the smaller this quantity is (i.e. perfect symmetry, like perfect repetition, in the output has a form measure of zero).

Measuring form at the output

A form measure is thus available for each test shape presented to the network. I.e., for the net trained on symmetry and the test set of random (neither symmetric nor repetitive) shapes, there were 20 different weight sets for the trained net, and 20 shapes to test, giving 400 shape-weight combinations. For each we can measure both the symmetry and the repetition of the shape generated by the net's output. Note that we are not concerned with the actual error involved, but with these measures that are based on the activation levels at the output.

Thus the "on-form" measure for a net shows the tendency of its output to resemble the general shape type with which it was trained,[2] with smaller quantities indicating greater affinity for that shape type. "Off-form" measures are also available (symmetry in the net trained on repetition, repetition in the net trained on symmetry) and were surprisingly important.

"On-Form" Measure

The general effect obtained in the model is striking. For the analysis described above, we find a significant interaction between shape type and step count when looking at the networks with respect to the type of stimulus used for training. This interaction can easily be perceived in the graphs through the much higher variance in the case of networks trained on repeated shapes (filled bars). Figure 4 shows the "on-form" analysis for the random test stimuli, comparing the degree of symmetry present in the symmetric nets, with the degree of repetition in the repetitive net. The interaction is highly significant: $F(2, 114) = 52.253, p < .001$).

This significant interaction between shape type and number of steps when we are using the measure appropriate for each network tells us one of two things. Either the networks, otherwise identical, have been differentially sensitized to step number by virtue of the type of shape they were trained on, or the manner in which activation is measured dictates that the quantity "tendency-toward-repetition" present in the output of the net will vary more than the the quantity "tendency-toward-symmetry," under the regime chosen to gauge it.

"Off-Form" Measure

We can clarify which of these is correct by examining the "off-form" measures. By looking at the output of the symmetrically trained net with "repetition goggles" and repetitively trained net with "symmetry goggles, " hopefully we can rule out the confound of this being a measurement effect.

In figure 5 a significant effect does obtain for the "off-form" measures of the random test stimuli ($F(2, 114) = 9.417, p < .001$). However, once more this is in the direction of repetitive nets showing more variance (by a factor of 2, upon examination of means). If the variance *was* due to the measurement of "tendency-to-repetition, " thus leading to a Type I error, we would now expect to see that variance in the symmetric when we are measuring *it* for its "tendency-to-repetition. " For the net trained on repetition, the "off-form" measure shows how symmetric its output is. The sustained effect in the net suggests a general sensitivity to step number, present even when we are examining how symmetric its output is.

[2] I.e. symmetric for the network trained on symmetry and repetition for the network trained on repeated shapes

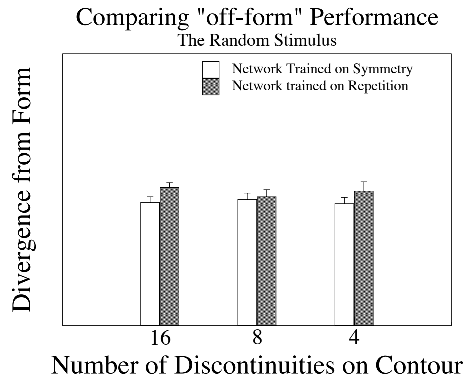

Figure 5: The compared "off-form" measures produced by trained networks, as a function of step number, under the *random* stimulus.

Discussion

The findings of the behavioral experiment were three fold, viz.

- the processing of shapes by the human observer varies qualitatively in accordance with the characteristics of shape's contours.
- the processing of symmetric shapes is carried out in parallel, while the processing of repetitive shapes remains a serial task, with point by point comparisons
- this is ecologically consistent with a cognitive facility that maximized correct figure ground segregation in a two dimensional image

The second of these points was behind the susceptibility to changes in step number on the contours of repetitive shapes only. Their data was consistent with the hypothesis, that processing of repetitive shapes is affected by the cognitive costs associated with serial processing, unlike the processing of symmetric shapes.

In our model, there is a similar distinction in the way the network handles these two classes of stimuli. In particular, when considering the degree to which the specific regularity of shape-type is learned in the network, we find generally better performance on symmetries, in contrast with repeated shapes. In addition, the performance of the nets trained on repetitive stimuli shows that they are affected by the number of discontinuities along the side of the pseudo-random block shapes. This is true for novel stimuli of the same type as well as for non-symmetric, non-repetitive block stimuli.

Interpretations: Step Number

The main effect that we would like to address here is the differential processing of symmetry and repetition by both human subjects and the split architecture.[3] The central finding was that there was a strong interaction between shape type and step number, with the repetitive shapes taking the brunt of the effect, far outstripping the variance induced by step number in their symmetric counterparts.

It certainly seems plausible that that difference is based on fixed anatomical features; looking for the the features held in common by both the model and the human subjects, the general split in the architecture of the processor is the obvious front-runner. This anatomical consideration in particular suggests how the differentiation of symmetry and repetition may be linked to the distinction between parallel and serial processing *through* the very structure of the processor.

The formalization of what constitutes repetition and symmetry in the block-shapes, activation across the vertical access summing to a constant in the former and having a net difference of zero in the latter, was crucial to the analysis presented here. But it goes further: it implies that for a split processor, symmetry is as simple as cross-checking (or cross-generating, in the case of this task) the output from each hidden layer. Repetition, on the other hand, involves a cumbersome re-calibration, for *every* segment of the repeated pattern, because the cross-image portion of the output is not a simple reflection.

Now if this difference is that symmetrical shapes can be checked by a parallel system, then it is conceivable that that system is rooted in the recognized image being split centrally along its axis of symmetry and each half being presented to each visual cortex, which in turn provides a massively parallel "cancellation" style verification of the image. Were we to assume homotopic and inhibitory commissures, symmetry would be exactly the reciprocal cancellation of activity across the two sides of the visual cortex, an idea that finds little favor in some circles (Dakin, 1998), but which a simple model such as the one here might help to refine.

But how does this explanation differ from that drawn in the original experiment? Baylis and Driver identified an aspect of shape processing in which symmetry was distinguished from repetition by rendering null the processing costs of increased shape complexity. This was elaborated as being a case of parallel versus serial processing, in which complexity, which is directly proportional to step number along the contour, only retarded the serial process, leaving the parallel process (the detection of symmetry) unhindered. Thus, the process of recognition is essentially one that involves the comparison of the segment end-points that make up the shape, and in the case of symmetry these comparisons take place simultaneously.

In some sense, the model is not incompatible with the take on the behavioral data. Indeed, the human study leaves open the question of what the facilitating mechanism is for the parallel treatment of symmetry, and the model provides one such possibility. Nevertheless, there is an important contrast. As discussed below, the notion of what constitutes complexity is not fixed. For Baylis and Driver, complexity increased with step number, and parallel processing was where such an increase was insignificant. For the model however, it would be quite an assumption to simply associate increased step number with increased complexity, for point-by-point comparisons of the stimuli have little meaning in a model lacking a temporal dimension. However, the difference between the symmetric and repetitive shapes is just as marked. In the absence of any sequential processing, this indicates that symmetry is special not in avoiding the narrow view of complexity-as-quantity, but in generally "playing-down" any dependence on contour variations. Since parallelism and seriality have no temporal meaning in the model there must be a more fundamental retreat from complexity that symmetry offers.

Complexity

Above, we alluded to the unfixed nature of complexity. A potential shortcoming in the modeling result is that the effect of step number seems to manifest in the reverse direction. That is, an increase in the number of steps in the stimuli presented to nets trained on repetitive shapes meant an increase in how well that nets output stayed to form (form in this case being repetitive). We would expect this "accuracy" to decrease, given that more steps presumably means greater complexity and therefore a harder task (and one for which, in the original experiment, the authors saw a need for more "counting" time). If anything, the results of the model disturb the clear relation obtaining between the original interpretation of serial versus parallel and how each accommodates effective increases in complexity.

Of course, there is no claim that the model attempts to perform either serial, or parallel processing, in the sense that Baylis and Driver use those terms. And it is hoped that the previous section went some way to reducing the high-level cognitive connotations of these terms to more concrete, anatomically based concerns. The networks learn to perform an identification task, and in doing so they pick up the general trends elicited by the stimuli sets used for training. It is in this sense then that processing of shape type differs: it will depend on how

[3] Though beyond the limits of this paper, it is worth noting that an effect of orientation found in the behavioral study, marked generally by poorer performance (RT and error), also falls out of the split-network model.

the task was learned in the first place.

In this context let us ask what complexity is. For though an increase in complexity can be equated with an increase in step number; for our model it won't be. Again, complexity in terms of the number of sequential operations performed (i.e. counting the steps) has no meaning in the context of the network. So how could a contour with 16 steps appear easier to process, or be in general more likely to encourage good "on-form" output, than one with only 8 or 4 discontinuities?

The answer to this involves reviewing the nature of the task, from the network's perspective: to reproduce accurately a contour of maximal discontinuity, which, in the case of this model, is one with 16 steps, the net at least has the advantage of avoiding any cross-row constraints. That is, given that the grain of the image and the grain of the shape in question match, no additional provisos are required in order for the net to attempt reproduction of the input at the output. Thus the task is relatively unencumbered, and the result is a more stable version of the form learned during training. But reduce the number of steps along the contour and the complexity of the task the net has to solve is actually *increased* by virtue of the added constraints of aligning rows of "pixels" at the output.

It isn't that this in itself disrupts the measure of form at the output, for that is always measured at the grain of the model, but that such additional constraints divert the resources that the net has devoted to producing well-formed images. The result is a drop in the form at the output, but, and this is the key point, this whole story, in which complexity for the network is revealed to be the opposite of what one would expect, only effects repetitive shapes.

Summary

The described model presents interesting analogues to some of the main effects uncovered in behavioral studies. In particular we have the preference for the processing of symmetric shapes, which is much less susceptible to variations along the contour than is the processing of repetitive shapes.

As already noted, symmetry may initially seem more complex a phenomenon, in terms of the operations required to generate symmetric contours (translation and reflection). However, from an anatomical perspective it may fact be simpler, especially around the vertical axis. This relates to the view that homotopicality between the visual cortices promotes the recognition of vertically oriented symmetry—because information, instead of being quantized and stored, can be "mirrored and checked" directly.

Baylis and Driver present a plausible argument for symmetry preferences in terms of parallel and serial processing, in the cognitive sense. However, more parsimonious explanations may be available. Using a split architecture neural net, we have suggested that the symmetry preference may arise from gross anatomical aspects of the processor. If this is so, then the application of Occam's Razor suggests that there is a simpler story on symmetry.

References

Baylis, G. C. & Driver, J. (1994) Parallel Computation of Symmetry but Not Repetition within Single Visual Shapes, *Visual Cognition*, **1** (4), 337-400.

Dakin, S. C., & Herbert, A. M. (1998) The Spatial Region of Integration for Visual Symmetry Detection. Proceedings of the Royal Society of London, B265, pp659-664

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. Cognition, 48, 71-99.

Hicks, J. (2002). A Computational Model of Split Processing. PhD Thesis (submitted), University of Edinburgh

Hicks, J. & Monaghan, P. (2001). Explorations of the interaction between split processing and stimulus types. In S. Wermter, J. Austin & D. Willshaw (Eds.) Emergent computational neural architectures based on neuroscience. Springer: Heidelberg.

Hicks, J., Oberlander, J. Shillcock, R.: Four letters good, six letters better: Exploring the exterior letters effect with a split architecture. Proceedings of the 22nd Annual Conference of the Cognitive Science Society. 2000. Madison, WI: Lawrence Erlbaum Associates.

Latecki, L. J. and Lakämper", Rolf (1999) Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution, *Computer Vision and Image Understanding: CVIU* V. 73, No. 3, 441–454

Sally, S., and Gurnsey, R., (2001) Symmetry detection across the visual field *Spatial Vision* Volume 14, No. 2, pp. 217–234,

Shillcock, R.C.,Monaghan, P. (2001) The computational exploration of visual word recognition in a split model. Neural Computation, 13, 1171-1198

G. J. ven Tonder, Y. Ejima, (2000) The patchwork engine: image segmentation from shape symmetries *Neural Networks* 13, 291-303