

Simulating Infants' Gaze Patterns during the Development of Perceptual Completion

Matthew Schlesinger

Department of Psychology
Brain and Cognitive Sciences
Southern Illinois University
Carbondale, IL 62901
matthews@siu.edu

Dima Amso

Sackler Institute for Developmental
Psychobiology, WMC
Cornell University
1300 York Ave., Box 140
New York, NY 10028
dia2004@med.cornell.edu

Scott P. Johnson

Department of Psychology
University of California,
Los Angeles
Los Angeles, CA 90095
sjohnson@psych.ucla.edu

Abstract

We have recently used a multi-channel, image-filtering model to study the development of visual selective attention in human infants (Schlesinger, Amso, & Johnson, 2007). In the current study, we employ the same model to simulate infants' gaze patterns during a perceptual completion task. The model not only succeeds in capturing the gaze patterns produced by young infants, but also provides support for the idea that a single neural mechanism may help to account for changes in visual selective attention and perceptual completion during early infancy.

1. Introduction

Learning to identify partly-occluded objects in a dynamic, 3-dimensional environment is a fundamental challenge for both natural and artificial visual systems. For human infants, a key element of this learning process is the development of *perceptual completion*, which involves the ability to integrate or link disjoint surfaces of an object that are separated by occlusion.

As we highlight in the next section, perceptual completion develops rapidly in infants during the first few months of life (e.g., Johnson, 2004; Slater, Johnson, Brown, & Badenoch, 1996). Several recent studies not only provide important clues about how this process unfolds, but they also suggest that visual activity and oculomotor skill play an essential role (e.g., Johnson, Slemmer, & Amso, 2004; Amso & Johnson, 2006).

We are investigating the hypothesis that the development of perceptual completion is made possible by advances in oculomotor skill. In particular, we have designed and tested a multi-channel, image-filtering model in order to identify potential neural circuits in the visual processing stream that may promote developmental changes in both visual selective attention and perceptual completion (Schlesinger, Amso, & Johnson, 2007).

To date, two critical findings have emerged from our modeling work. First, the model is able to account for progressive improvements in infants' visual selective attention as measured by performance on a visual search task. Second, and more importantly, the model has implicated a specific cortical area—the posterior parietal

cortex—as a neural substrate whose development supports efficiency in visual selective attention and in turn perceptual completion.

Our current goal is to demonstrate that the same model can also account for the development of perceptual completion. We investigate this claim by analyzing the model's simulated eye movements during the unity perception task, which is used to study the development of perceptual completion in young infants.

The rest of the paper is organized as follows. Section 2 briefly reviews the method and findings from the unity perception task. In particular, we highlight the evidence for a link in early infancy between visual selective attention and perceptual completion. In Section 3, we provide an overview of the image-filtering model, and illustrate how it is used to simulate infants' eye movements. Section 4 presents the results from a simulation study that replicates the gaze-shift patterns produced by human infants during the unity perception task. In the final section, we provide an integrative account of our findings, while also highlighting some of the model's limitations.

2. Perceptual completion in infants

A key technique for studying the development of perceptual completion in human infants is the *unity perception task*. In this paradigm, infants first view a solid rod that moves laterally while its center is occluded by a large screen (see Figure 1A, “occluded rod”). This display is repeated until infants lose interest in it (i.e., they become habituated). Infants are then shown two test displays in alternation. In one display (Figure 1B, “complete rod”), a single rod with no occluding screen is presented, while in the other (Figure 1C, “broken rod”) two rod segments (matching the upper and lower portions

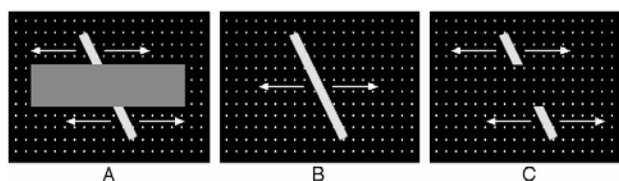


Figure 1: Displays used to assess perceptual completion in infants: (A) occluded-rod (habituation) display, and (B) complete-rod and (C) broken-rod test displays.

from the habituation display) are presented.

The tendency to look longer at one test display over the other is assumed to reflect a novelty preference (e.g., Bornstein, 1985), and in particular, can be used to infer that infants at a given age perceive the occluded rod as either (a) a single, center-occluded object, or (b) two separate rod segments that move synchronously.

Indeed, newborn infants look significantly longer at the complete rod during the test phase (e.g., Slater, Johnson, Brown, & Badenoch, 1996). This pattern suggests that neonates perceive the occluded rod as two disjoint surfaces or objects, and therefore find the complete rod to be novel or unfamiliar. In contrast, 4-month-olds look significantly longer at the broken-rod display, suggesting that infants at this age perceive the occluded rod as a single, unitary object (i.e., *perceptual unity*; e.g., Johnson, 2004).

2.1 Visual selective attention and perceptual completion

This developmental pattern raises an important question: if perceptual completion is not present at birth, what postnatal experiences or abilities (if any) are necessary for it to develop? Amso and Johnson (2006) proposed that perceptual completion may depend on oculomotor skill, and in particular, on the development of *visual selective attention*. Visual selective attention includes the ability to deploy attention in a deliberate or goal-directed manner, while ignoring irrelevant or distracting stimuli.

Amso and Johnson (2006) hypothesized that if visual selective attention is a prerequisite for perceptual completion, then separate measures of each capacity should be closely related in early infancy. This reasoning was used to motivate two predictions. First, they predicted that infants who perceive unity—based on their posthabituation looking-time preference—should scan the occluded-rod display differently than infants who have not yet developed unity perception. In particular, infants who perceive unity should focus more attention on the rod segments and the motion of the rod than infants who do not perceive unity.

To address this prediction, Amso and Johnson (2006) recorded both the eye movements and global looking times of 3-month-olds during the unity perception task. Infants were then sorted into one of two groups as a function of their posthabituation looking times. Half of the infants were classified as *perceivers*, as they looked longer at the broken-rod test display, while the other half did not look longer at either test display, and were therefore classified as *non-perceivers*.

Next, perceivers' and non-perceivers' eye-movement patterns during habituation to the occluded-rod display were compared. *Rod scans* were defined as the proportion of saccades (i.e., gaze shifts) that either began and ended on one rod segment, or travelled from one rod segment to the other. Meanwhile *vertical scans* were defined as saccades that travelled from a lower to a higher quadrant, or vice versa, without contacting the rod. Amso and Johnson (2006) reasoned that rod scans should be positively associated with unity perception, while vertical

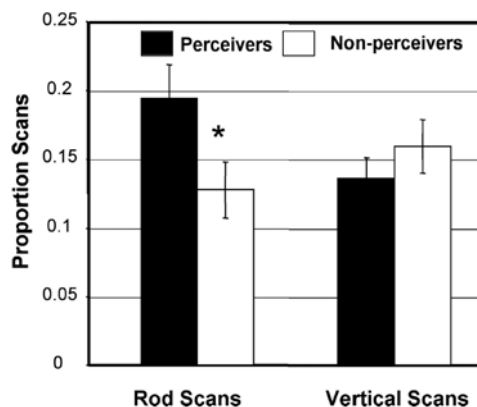


Figure 2: Proportion of rod scans and vertical scans in 3-month-old perceivers vs. non-perceivers (in Amso & Johnson, 2006). Error bars indicate ± 1 SEM. * $p < .05$

scans were not expected to be diagnostic (i.e., to differentiate perceivers from non-perceivers).

As Figure 2 illustrates, perceivers generated a significantly higher proportion of rod scans than non-perceivers. Meanwhile, there was not a significant difference between either group of infants in their proportion of vertical scans. Thus, perceivers appear to tailor their scanning patterns toward the occluded-rod display more effectively than non-perceivers, and in particular, perceivers focus more attention on the features of the display that are essential for forming the percept of a single, integrated rod (for related findings, see also Johnson, Slemmer, & Amso, 2004).

Second, Amso and Johnson (2006) also reasoned that if infants' performance on the unity perception task is due to underlying differences in visual selective attention, then perceivers should be more successful than non-perceivers on a visual search task. Consequently, the same 3-month-olds described above were also presented with a series of visual "pop-out" displays, in which a tilted bar was embedded within a field of vertical bars. As predicted, infants who had been classified as perceivers succeeded in detecting the tilted target significantly more often than those who were classified as non-perceivers.

2.2 Modeling the development of visual selective attention

Taken together, these findings are consistent with the view that visual selective attention serves to guide the development of perceptual completion. One way to account for this link is to propose that development within a single, underlying neural pathway supports changes in both visual selective attention and perceptual completion.

Schlesinger, Amso, & Johnson (2007) pursued this account with the use of a computational model of early visual processing. The goal of the simulation study was to identify three candidate neural systems, and to hand-tune corresponding parameters in the model while simulating infants' performance on the visual search task studied by Amso and Johnson (2006).

One of the candidate systems chosen for study was the

posterior parietal cortex, based on findings that this area of the visual system is involved in the encoding of stimulus salience, as well as the production of overt and covert attentional shifts (e.g., Gottlieb, Kusunoki, & Goldberg, 1998; Itti & Koch, 2000). As we highlight in the next section, recurrent (or self-sustaining) activity in the posterior parietal cortex is represented in the model by a parameter that can be systematically increased or decreased, resulting in a longer or shorter duration of recurrent activity.

A key finding from the simulation study was that the model successfully captured the real-time visual search performance (i.e., detection rates and search latencies) of perceivers and non-perceivers. In particular, the model transitioned from the performance level of non-perceivers to that of perceivers as the duration of recurrent parietal processing was increased.

An important next step in evaluating the model is to determine whether it can also account for the gaze patterns produced by perceivers and non-perceivers in the unity perception task (i.e., during the occluded-rod display). In the next section, we provide a detailed description of the model, as well as how it is used to simulate infants' gaze patterns.

3. The image-filtering model

The image-filtering model is designed to simulate the sequence of transformations that occur as the optic array is projected onto the retina, and then transduced through a series of neural structures that function like linear image filters (e.g., lateral geniculate nucleus, visual cortex, parietal cortex, etc.). The model not only captures several functional properties of the mammalian visual system (e.g., center-surround receptive fields, lateral or horizontal inhibitory connections in area V1, etc.), but it is also consistent with multi-stage computational theories of vision that describe perceptual processing in terms of a feature-extraction stage, followed by a feature-integration or combination stage (e.g., Kastner & Ungerleider, 2000; Treisman & Gelade, 1980).

The model was originally developed by Itti and Koch (2000), but has been adapted here for the purposes of simulating the production of overt and covert attentional shifts in young infants. In particular, three specific modifications include (1) the addition of a motion channel during the feature-extraction stage, (2) the elimination of recurrent processing in the four conspicuity maps, and (3) the use of a stochastic decision-rule for producing gaze shifts (each of these features is described below).

Figure 3 illustrates the processing of a two-dimensional visual input by the model into a sequence of retinotopic maps. As the figure highlights, the input image is first “decomposed” across four parallel feature channels (i.e., intensity, motion, color, and orientation; see Figure 3B). Next, individual feature maps are combined into conspicuity maps (see Figure 3C), and then these maps are summed into a salience map (see Figure 3D).

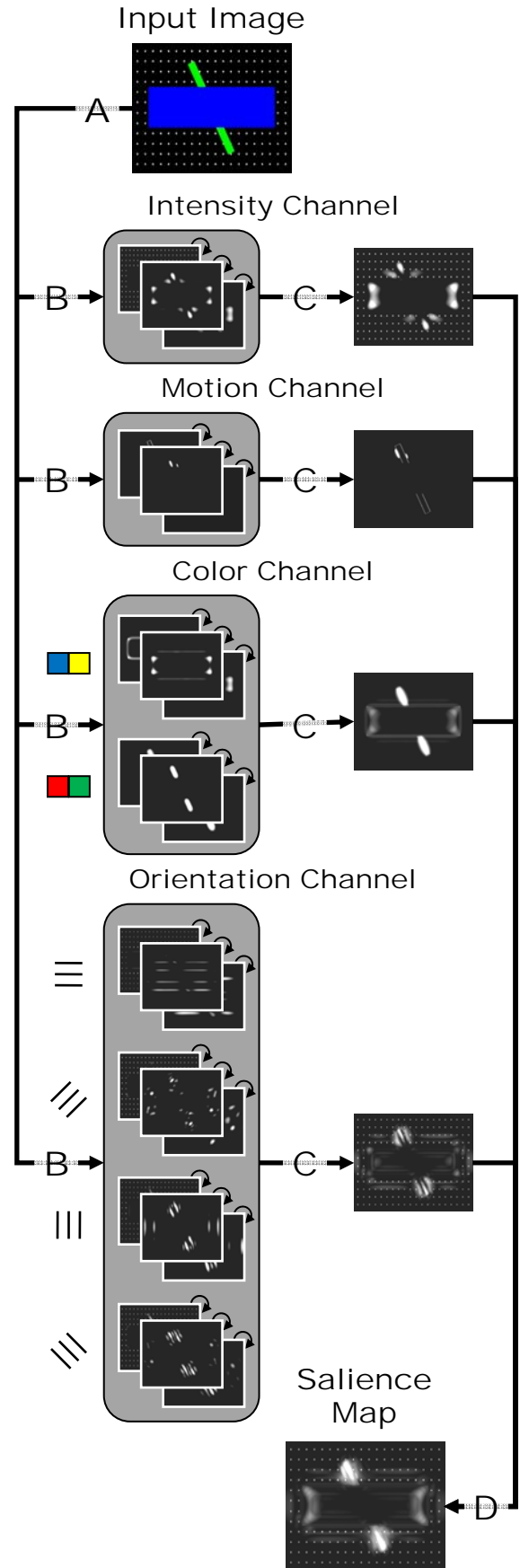


Figure 3: A schematic diagram of the image-filtering model. Processing stages include propagation of the input image (A), feature extraction, center-surround contrast, and feature competition (B), feature pooling into conspicuity maps (C), and summing into a salience map (D).

3.1 Description of the model

We provide here a brief description of each of the major processing stages in the image-filtering model, while additional details are reported in Schlesinger, Amso, & Johnson (2007). A complete mathematical description of each of the feature-extraction algorithms is presented by Itti, Koch, & Niebur (1998).

(1) Feature extraction. The initial stage of processing in the image-filtering model represents the sequence of transformations that occur as light is transduced into electrical signals on the retina, and then propagated from the retinal ganglion cells, through the lateral geniculate nucleus, to the visual receiving area in occipital cortex.

As Figure 3B illustrates, the model includes 8 total features, divided into four feature channels or dimensions: 1 intensity feature (i.e., brightness), 1 motion feature, 2 color features (i.e., blue-yellow and red-green opponent pairs) and 4 orientation features (i.e., 0°, 45°, 90°, and 135°). Each feature is encoded at three spatial scales (i.e., fine, medium, and coarse), resulting in a total of 24 feature maps. Note that for each set of 3 maps in Figure 3B, the top map that is displayed corresponds to the medium spatial scale.

(2) Center-surround processing. Figure 3B also includes the second stage of processing, in which a center-surround receptive-field filter is applied to each of the 24 feature maps. This processing stage not only mimics the inhibitory-excitatory organization found in early visual processing, but in computational terms, it also serves to reduce noise while increasing feature contrast in each map.

(3) Within-feature competition. Note that each feature map in Figure 3B has a curved arrow in the upper-right corner. This arrow represents the third stage of processing in the model, where a local-excitation/global-inhibition filter is applied to each feature map.

Two aspects of the feature-competition filter are particularly important. First, the filter employs a recursive or iterative process, in which the output of the filtering operation becomes the input on the next iteration of the process. Second, while the pattern of activity over a feature map tends to vary from iteration to iteration, in general a stable pattern typically emerges after several iterations of the filtering operation (e.g., 9 or 10 iterations or loops). In effect, the feature-competition filter suppresses similar features on a map while increasing the strength of unique or distinct features (see Figure 7).

In neurophysiological terms, the feature-competition filter corresponds to covert attentional processing in the posterior parietal cortex (e.g., intraparietal sulcus, Gottlieb, Kusunoki, & Goldberg, 1998). In particular, recurrent or self-sustaining patterns of neural activity in parietal cortex enable two or more salient locations in the visual field to “compete” for attention through mutual inhibition. Thus, the number of discrete “parietal” loops in the model is intended to represent the temporal duration of recurrent activity in parietal cortex.

(4) Conspicuity and salience maps. After the within-feature competition has terminated, the feature maps within each feature channel are summed into a

conspicuity map that represents the corresponding channel (see Figure 3C). Next, the four conspicuity maps are summed, creating a unified salience map (see Figure 3D).

(5) Target selection. The final processing stage involves selecting a target from the salience map for fixation. Itti and Koch (2000) employ a deterministic selection strategy, in which the most salient location in the visual field is fixated first, followed by the second most salient location, and so on. However, such a strategy may fail to capture two important features of oculomotor control.

First, the phenomenon of inhibition-of-return (IOR), observed in both infants and adults, is the tendency to avoid returning to a recently-fixated location (e.g., Johnson, 1994; Hood, 1995). Second, when two targets are equally salient, there may also be a tendency to prefer the closer target (i.e., “proximity preference”; see Koch & Ullman, 1985).

In order to incorporate these constraints into the target-selection process, two adjustments to the salience map are performed prior to each gaze shift. First, a location-inhibition map is updated synchronously with the salience map (the fixation map is not shown in Figure 3). This inhibition map maintains a history of the salience level at each of the locations previously fixated by the model. When added to the salience map, the location-inhibition map lowers the salience at those locations. Note that activity on the location-inhibition map decays over time, in order to allow the model to eventually return to previously-fixated locations.

Second, after the fixation and salience maps are combined, activation at each location on the resulting map is divided by the distance to that location from the current fixation point. As noted above, this adjustment incorporates a minimum-distance bias on the production of saccades.

As in our previous model, we employ a stochastic process for selecting each fixation target. Specifically, a *softmax* function is used to assign a probability of fixation to the 100 most active locations on the adjusted salience map. This function not only introduces variability into the selection process, but also ensures that locations with moderate (but not high) salience are occasionally selected.

3.2 Evaluating the model

In order to evaluate the model’s eye-movement patterns during the occluded-rod display, the animation stimulus used by Amso and Johnson (2006) was parsed into still frames, and presented to the model one frame at a time. The duration of the occluded-rod display is 5 seconds with a framerate of 30fps, which resulted in 150 still frames, each with a dimension of 480x360 pixels. Following the sequence of computational stages described above, each still frame was processed by the model and generated a corresponding pattern of activity on the salience map (see Figure 3).

Note that the model did not produce a saccade on each time step, as this would result in an unrealistically high saccade frequency (i.e., 30 eye movements per second!). Instead, a threshold rule was used to determine the

saccade latency (i.e., the interval between saccades), by sampling from a normal distribution with a mean of 200ms and standard deviation of 20ms. These values were derived from Amso and Johnson's (2006) estimate of infants' mean saccade frequency while they viewed the occluded-rod display (i.e., $M = 4.57$ saccades per second, $SD = 0.52$).

When the temporal threshold was reached (i.e., a given input image was selected to produce a saccade), the salience map was computed and adjusted as described above, and a simulated gaze shift was performed. For time steps that were not selected for a saccade, activation on the salience map continued to accumulate over successive input images, until the next saccade was produced. Note that after each saccade, activity values for all locations on the salience map were reset to 0.

A single simulation run was designed to correspond to a testing session with one infant. At the start of a run, the model's virtual fixation point (i.e., point of gaze) was positioned at a random location on the first input image, and then a series of saccades was produced while the remaining images in the animation sequence were presented to the model. In order to generate a sufficient sample of saccades from each simulated infant, the sequence of images from the occluded-rod display was repeated for 20 cycles (i.e., 3000 frames or 100 seconds).

The testing procedure began with the number of parietal loops set at 0 (i.e., the feature-competition filter was not applied). Ten simulated subjects were run at this setting. Next, the number of loops was incremented from 0 to 1, and the testing procedure was repeated. This process continued until the number of parietal loops reached 10.

Figure 4 illustrates the six areas of interest (AOIs) used to code the model's saccades. Areas 1 and 2 correspond to the upper and lower segments of the rod, respectively, while the remaining four areas correspond to the four quadrants. Note that a mutual-exclusivity constraint applied, such that each fixation could only be coded as contacting a single AOI. In particular, fixations to either of the rod AOIs took precedence over the four quadrants.

Rod scans were defined as a saccade in which the start and end point of the gaze shift either (a) remained within one of the rod AOIs (i.e., a lateral saccade, following the movement of a rod segment), or travelled from one rod AOI to the other. Meanwhile, *vertical scans* were defined

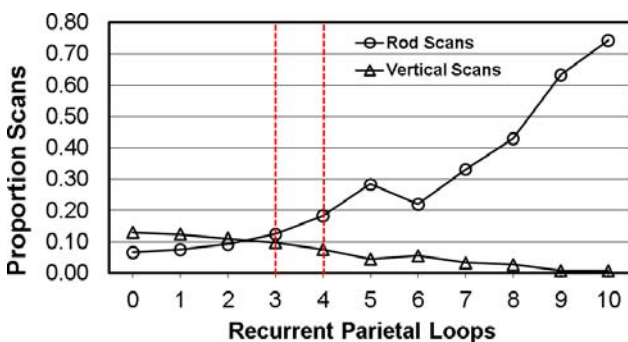


Figure 5: Simulated proportion of rod scans (circles) and vertical scans (triangles) produced by the model, as a function of the number of recurrent parietal loops (see text for details).

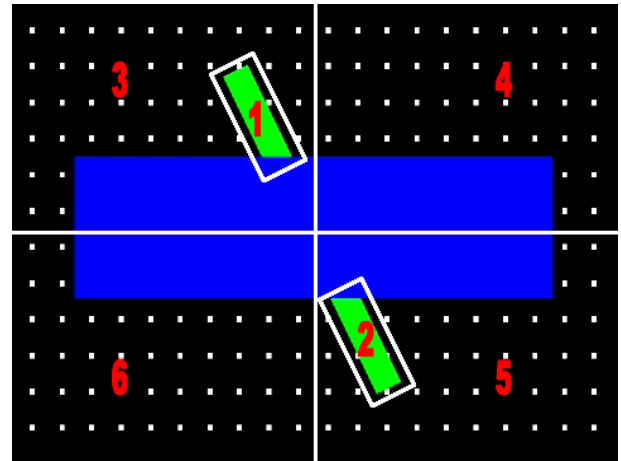


Figure 4: Six areas of interest (AOIs) used for coding saccades in the image-filtering model.

as a saccade from a lower to an upper quadrant, or vice versa, without contacting either of the rod AOIs.

4. Simulation results

Figure 5 presents the proportion of rod and vertical scans produced by the model, respectively, as a function of the number of parietal loops. As the model indicates, the proportion of rod scans increases as the number of recurrent parietal loops is increased, with the exception of the transition from 5 to 6 loops. While it is plausible that this overall trend may represent a general developmental pattern—that is, a higher proportion of rod scans with age—it also seems likely that the upper end of the figure (e.g., from 8 to 10 parietal loops) over-estimates the proportion of rod scans that older infants may produce.

Figure 5 also indicates that the proportion of vertical scans is negatively associated with the number of parietal loops. This trend is due in part to the fact that the total number of saccades produced by the model is fixed, so that as rod scans increase, the proportion of other saccade patterns must necessarily decrease.

Recall that Amso & Johnson (2006) found that perceivers generated more rod scans than non-perceivers. More precisely, as Figure 2 illustrates, about 20% of perceiver's saccades were rod scans, while the percent of rod scans produced by non-perceivers was roughly 13%. The two vertical, dashed red lines shown in Figure 5

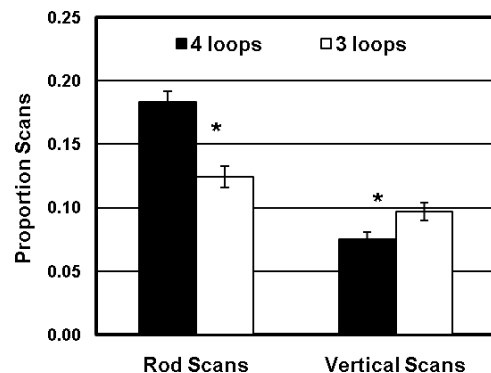


Figure 6: Simulated proportion of rod scans and vertical scans in the model at 3 and 4 parietal loops. Error bars indicate ± 1 SEM. $*p < .01$

indicate that the model’s performance corresponds to that produced by perceivers and non-perceivers when the number of parietal loops is set at 4 and 3, respectively.

In order to provide a relatively direct comparison between the model and 3-month-old infants, Figure 6 presents a subset of the data from Figure 5. In particular, Figure 6 presents the proportion of rod scans and vertical scans produced by the model when the number of recurrent parietal loops is set at 3 versus 4. As the figure highlights, the performance pattern at 3 loops corresponds to the pattern produced by non-perceivers, while increasing the number of loops in the model to 4 results in a performance pattern that is comparable to that generated by perceivers.

While there are several qualitative similarities between the model’s and infants’ gaze patterns, two differences should be noted. First, the model appears to generate a lower overall proportion of vertical scans than 3-month-olds. Second, while the difference in vertical scans between 3-month-old perceivers and non-perceivers was not significant, the corresponding difference was significant in the model. One way to account for these differences between the model and human infants may be to more closely examine variability between subjects, as it is likely that infants are more variable in their gaze behavior than the model.

Before concluding this section, we offer a snapshot of the model’s internal state during processing, in order to help the reader visualize the functional consequence of increasing the number of recurrent parietal loops—or in computational terms, extending the duration of the within-feature competition stage.

Figure 7 illustrates the state of the salience map in response to input image 39, as the number of recurrent parietal loops or iterations is incremented from 0 to 10. In contrast to Figure 3, we present the salience maps here as 3D topological surfaces, to highlight changes in activation (i.e., the height of features on the map) across iterations.

First, note that during iteration 0, not only the rod segments and vertical edges of the box are active on the salience map, but also the horizontal edges of the box as well as the background texture elements. This suggests that the feature-extraction filters succeed in capturing a wide array of sensory features in the occluded-rod display. However, after a single iteration of the within-feature competition stage (i.e., iteration 1), much of the “background” activity on the salience map has diminished. This is the result, of course, of higher activations at the rod segments and the vertical edges of box suppressing or inhibiting activity at other locations on the salience map. In quasi-psychological terms, we might interpret this change as corresponding to an increase in the perceptual “focus” of covert attention.

Second, note also that activity at the upper and lower rod segments continues to strengthen over iterations, as activity at the vertical edges of the box diminishes. This pattern helps to explain the gradual increase in the proportion of rod scans produced by the model as recurrent parietal loops is increased (see Figure 5), since locations with higher activity are more likely to be fixated by the model.

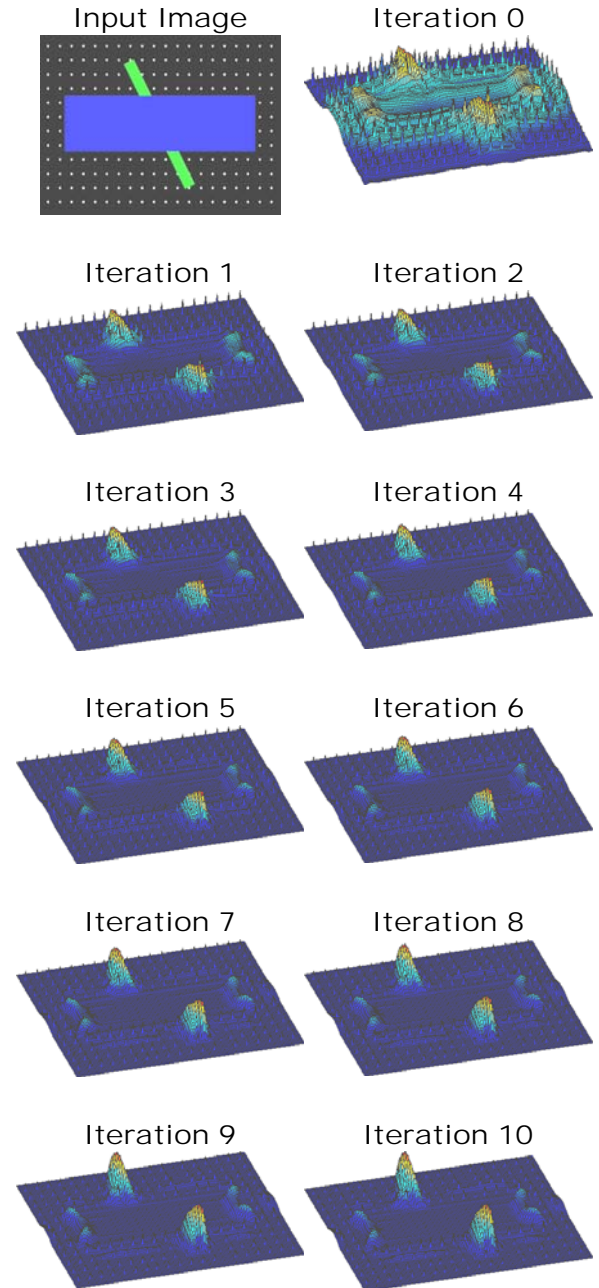


Figure 7: A sample input image (frame #39), and the resulting activation state of the salience map, over 10 iterations of the recurrent parietal loop.

5. Discussion

Recent behavioral research with infants has identified an important connection between perceptual completion and visual selective attention (Amso & Johnson, 2006). This work is consistent with the idea that the ability to deploy visual attention efficiently and effectively helps to support the development of perceptual completion—that is, the percept that disjoint surfaces belong to a single, unified object.

We have proposed and tested a computational model of early visual processing, in order to investigate the idea that one underlying neural circuit may provide a foundation for both the development of perceptual

completion and visual selective attention. As a first step, our prior simulation work focused on visual selective attention (Schlesinger, Amso, & Johnson, 2007). In particular, the model captured one of two major findings reported by Amso & Johnson (2006): by increasing the duration of recurrent parietal processing, the model replicates the shift in visual search performance from non-perceivers to perceivers.

Amso & Johnson's (2006) second major finding is that perceivers are also more effective at scanning the occluded-rod display than non-perceivers. Is this difference because perceivers are more focused scanners, and less distracted by extraneous visual stimuli?

We addressed this question in the current simulation study by again varying the duration of recurrent parietal processing in the model, while measuring the resulting gaze-shift patterns produced during the occluded-rod display. We indeed found not only that (a) the model captures the specific gaze-shift patterns produced by perceivers and non-perceivers, but also more generally, (b) the model produced more rod scans as recurrent parietal feedback was increased.

Our findings from the two simulation studies suggest a series of tentative conclusions. First, a key feature of early processing in the mammalian visual system is within-feature competition (e.g., Gilbert & Wiesel, 1989; Sillito & Jones, 1996). As Figure 7 illustrates, this competitive process tends to enhance the salience of one or a few locations in the visual field, while progressively lowering the influence of stimuli at other locations. Therefore, such a mechanism may help to account for the difference between perceivers and non-perceivers in terms of visual selective attention.

Second, following this line of reasoning, we might also conclude that more effective and efficient scanning strategies emerge in young infants as recurrent parietal processing progressively develops. In other words, as feedback mechanisms enable increasingly longer durations of within-feature competition, infants should appear more selective about how they deploy their attention over space and time.

Third, with more efficient and effective scanning, infants may also be more likely to detect or discover relevant perceptual regularities or patterns (e.g., Gestalt principles such as "common fate"), and to incorporate those patterns into their visual activity (e.g., Cohen, Chaput, & Cashon, 2002).

This developmental account remains somewhat preliminary, however, until a number of important issues have been addressed. In particular, the current model can be elaborated in at least two ways. First, note that in both of our simulation studies, we explicitly hand-tuned or modulated the number of recurrent parietal loops. While this change is intended to serve as a proxy for a developmental process, it remains to be demonstrated that such a change can occur adaptively (i.e., as the result of an autonomous learning process). We are currently exploring this question with the use of a reinforcement-learning algorithm.

Second, it should also be noted that while the model successfully captures the eye-movement patterns generated by both perceivers and non-perceivers, it is

likely that the internal state (i.e., representations) produced by the model differ in one key respect from those of perceivers, in particular. That is, the phenomenal experience of perceptual completion may be that one can easily imagine, or "virtually" see the occluded portion of a partially-visible object. Supporting this phenomenology are neurons in early visual areas (e.g., V2) that literally fill-in the occluded space between two segments of an object (e.g., Albright & Stoner, 2002). We are also interested in incorporating a comparable neural mechanism into the image-filtering model, and investigating its effect on the model's eye-movement patterns.

In conclusion, we note that the concept of *active vision* and the progressive construction of a more complex and detailed visual world are themes shared by researchers in both psychology and robotics. In addition, we suspect that the developmental pattern we are investigating is not limited to perceptual completion, but may be a ubiquitous phenomenon throughout early perceptual and cognitive development (e.g., object recognition, object permanence, etc.).

References

- Albright, T.D., & Stoner, G.R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, 25, 339-379.
- Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, 42, 1236-1245.
- Bornstein, M.H. (1985). Habituation of attention as a measure of visual information processing in human infants: Summary, systemization, and synthesis. In G. Gottlieb & N.A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 253-300). Norwood, NJ: Ablex.
- Canfield, R.L., & Kirkham, N.Z. (2001). Infant cortical development and the prospective control of saccadic eye movements. *Infancy*, 2, 197-211.
- Cohen, L.B., Chaput, H.H., & Cashon, C.H. (2002). A constructivist model of infant cognitive development. *Cognitive Development*, 17, 1323-1343.
- Gilbert, C. D., & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience*, 3, 1116-1133.
- Gottlieb, J.P., Kusunoki, M., & Goldberg, M.E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391, 481-484.
- Hood, B. M. (1995). Shifts of visual attention in the human infant: A neuroscientific approach. In C.Rover-Collier, & L. Lipsett (Eds.), *Advances in infancy research* (pp. 163-216). New Jersey: Ablex.

- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489-1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual-attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254-1259.
- Johnson, M. H. (1994). Visual attention and the control of eye movements in early infancy. In C. Umiltà, & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 291-310). Cambridge, MA: MIT Press.
- Johnson, S.P. (2004). Development of perceptual completion in infancy. *Psychological Science*, 15, 769-775.
- Johnson, S.P., Slemmer, J.A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, 6, 185-201.
- Kastner, S., & Ungerleider, L.G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, 23, 315-341.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4, 219-227.
- Schlesinger, M., Amso, D., & Johnson, S.P. (2007). The neural basis for visual selective attention in young infants: A computational account. *Adaptive Behavior*, 15, 135-148.
- Sillito, A. M., & Jones, H. E. (1996). Context-dependent interactions and visual processing in V1. *Journal of Physiology - Paris*, 90, 205-209.
- Slater, A., Johnson, S.P., Brown, E., & Badenoche, M. (1996). Newborn infants' perception of partly occluded objects. *Infant Behavior and Development*, 19, 145-148.
- Treisman A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.