Schlesinger, M.; Amso, D.; Johnson, S.P.; Hantehzadeh, N.; Gupta, L., "Using the iCub simulator to study perceptual development: A case study," Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on Development and Learning, pp.1,6, 7-9 Nov. 2012 doi: 10.1109/DevLrn.2012.6400866

Using the iCub Simulator to Study Perceptual Development: A Case Study

Matthew Schlesinger, Dima Amso, Scott P. Johnson, Neda Hantehzadeh, and Lalit Gupta

***Abstract*—In the present study, we assess the iCub simulator as a platform for investigating perceptual development in human infants. In particular, we evaluate the simulator as a quasi-realistic virtual environment for conducting object-perception experiments. As a case study, we describe our simulation of the *perceptual-completion task*, which measures infants' perception of a moving, partially-occluded object. We present here two simulation studies. The first study replicates our previous findings, which demonstrate that increasing spatial competition in our eye-movement model results in increased attention toward the occluded object. We then extend these findings in a second simulation study by demonstrating that reducing the width of the occluder also increases attention toward the occluded object, though unexpectedly, only for low levels of spatial competition. We conclude by highlighting the value of the iCub simulator as a research tool for psychologists, and also by noting how our eye-movement model can be further improved and elaborated.**

***Index Terms*—perceptual development, iCub robot simulator, object perception, spatial competition.**

I. Introduction

THE emerging field of developmental robotics provides researchers from the machine-learning and human-behavior communities with the means to work collaboratively on the study of development in natural and artificial systems. In particular, both communities mutually benefit by adopting a shared set of theories, methodological tools, and experimental paradigms [1]-[2]. The purpose of the current paper is to highlight a specific tool-kit, designed within the domain of robotics – that is, an immersive, quasi-realistic robot simulator – which provides researchers with a platform for modeling and investigating perceptual development in human infants.

We have three objectives. First, we argue that the use of an immersive, embodied simulator is a logical next step in the process of designing and testing computational models of development. In particular, we briefly survey the range of modeling approaches that have been proposed over the last two decades, and demonstrate that subsequent models have narrowed the gap between the study of perception in human infants, and the study of corresponding skills or abilities in computational models.

Second, as a specific case study, we illustrate our approach by presenting the perceptual-completion task, which is used to investigate infants' perception of a moving, partially-occluded object. In this section, we also describe our eye-movement model [3]-[5]. The model is designed to simulate the gaze behavior of an infant within an experimental paradigm like the perceptual-completion task, and it provides a mechanistic account of how changes in the neural substrates of visual attention can influence the process of perceptual development.

Lastly, we present the results from two simulation studies, in which our model is embedded and tested within the iCub simulator. The first study replicates the findings from our previous modeling framework, which represents the infant as a minimally-embodied agent. By testing our model in the iCub simulator – which adds a considerable degree of physical embodiment and realism – our successful replication results provide an important check on the robustness and validity of the model. We then describe a second study that extends our findings by presenting the simulated infant with a modified version of the perceptual-completion task.

II. An Overview of Modeling Approaches

In this section, we highlight four classes of computational models of perceptual and cognitive development that have been proposed over the last 20 years: neural-circuit models, retinal-input models, active-vision models, and embodied-robotic models.

A. Neural-Circuit Models

The first class of developmental models was heavily influenced by the success of connectionism in the late 1980s, and as a result, these models typically employed a three-layer neural-network architecture, with learning via the backpropagation-of-error rule. We refer to these models as *neural-circuit* models, as they are conceptualized as embedded within the agent's nervous system and "distant" from the environment in at least two ways: first, input to the network is often pre-processed into predefined features, and second, the

Manuscript submitted June 29, 2012.

M. Schlesinger is in the Department of Psychology at Southern Illinois University Carbondale, Carbondale, IL 62901 USA (phone: 618-453-3524; fax: 618-453-3563; e-mail: matthews@siu.edu).

D. Amso is with the Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI 02912 USA (email: Dima_Amso@brown.edu).

S. P. Johnson is with the Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095 USA (email: scott.johnson@ucla.edu).

N. Hantezadeh and L. Gupta are with the Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, Carbondale, IL 62901 USA (e-mail: neda@siu.edu and lgupta@siu.edu).

output of the network is an abstract "choice" or "decision," which is interpreted by the model-builder as representing a particular action, but is not overtly performed by the simulated infant (e.g., a motor signal that is not converted into an actual movement).

An example of this approach is illustrated by [6], who simulate the development of perceptual categorization in young infants. In particular, an autoencoder network is first trained to represent a set of dogs, and then the network is tested on a set of cats (or vice versa). Rather than presenting the network with images of dogs and cats, however, each image is decomposed into a set of basic features (e.g., head size, nose length, etc.), and presented to the network as a vector of feature values (i.e., a localist encoding). A similar modeling strategy is also employed by [7], which simulates infants' search behaviors for hidden objects. In addition to preprocessing the input into a feature-vector, in this model the output of the network (i.e., the search behavior) is represented by a pattern of activity over a set of units that encode specific locations; the output unit with the highest activation level is interpreted as the place where the model chooses to "look for the hidden object."

B. *Retinal-Input Models*

Subsequent models refined the neural-circuit approach by reformulating the input to the network, so that the input layer makes "contact" with the environment through a retina-like input representation. We therefore refer to these as *retinal-input* models. A common strategy for retinal-input models is to structure the output layer of the network with the same number and arrangement of units as the input layer, and then to either teach the network to reproduce the input pattern on the output layer (i.e., an autoencoder network), or to produce the expected next input pattern (i.e., a prediction network).

Examples of retinal-input models include [8] and [9], which both simulate the capacity to track moving objects that are briefly occluded. In addition, a more recent retinal-input model that relies on prediction-learning is [10]. An important and impressive feature of [10] is that it demonstrates that the same architecture and learning algorithm can account for developmental changes in two related perceptual abilities (i.e., perceptual completion and tracking of occluded objects). An alternative model of perceptual completion that adopts a hybrid strategy is proposed by [11], in which both preprocessing of basic visual features and a retinal-input representation are implemented.

C. *Active-Vision Models*

Active-vision models provide an additional step toward an embodied and situated agent by modifying the function of the output produced by the model. In particular, the output of the conventional retinal-input model is reconfigured so that it enables the simulated infant to produce discrete eye movements. An early example of this approach is [12], who simulate a moving "fovea" that learns to track both visible and occluded objects. In a subsequent model, [13] simulates a

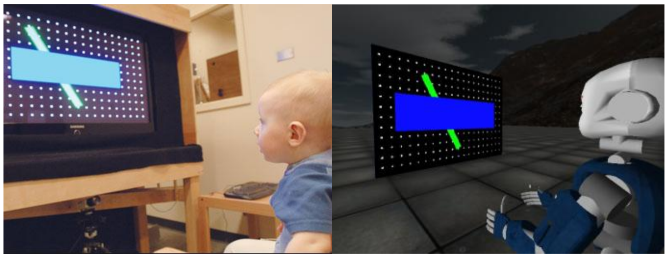

Fig. 1. An infant views the occluded-rod display during the perceptual-completion task (left), and the iCub robot views a corresponding display in the iCub simulator (right).

dual-channel visual input system that receives input through both a low-resolution peripheral channel and a high-resolution fovea channel. Like [12], the simulated infant in [13] responds by moving its fovea along a 2D plane. Another important feature of [13] is that rather than presenting the model with computer-generated input images, a live scene was videorecorded, digitized, and used to train and test the model.

D. *Embodied-Robotic Models*

An important limitation of the active-vision models that have been studied thus far is that they are minimally-embodied. In our own work, for example, the infant's "body" consists of a visual receptor surface (i.e., retina), which projects to a series of visual processing areas, which activate a highly-simplified motor system that generates eye movements ([3]-[5], [12]-[13]). It should be noted that, besides the fact that the simulated infant has no physical extent in 3D space, the visual input is projected directly on to the retina. Thus, the geometric properties of the environment (e.g., linear perspective) are not well-represented by the active-vision model, and spatial cues that might play an important role during eye-movements (e.g., motion parallax) are unavailable.

A logical next step in the design and testing of developmental models, which not only addresses these limitations, but also provides a number of additional advantages, is to embed the model within a robotic agent that is situated in a real or simulated environment. We briefly describe here a few of the benefits:

1) Robotic models have direct access to the environment through their sensors, and can use this sensory data to detect and represent potentially important geometric and/or physical cues and features. This greatly simplifies the design question of "how input to the model occurs," by allowing the simulated infant in the robotic model to "read off" the state of the environment via direct sensation.

2) As Figure 1 illustrates, the environment of the robotic model can be designed to emulate the context in which the human infant is tested, including the appearance of the environment and the stimuli used to test the simulated infant. By increasing the similarities between the model and the infant experiment, the robustness, validity, and explanatory power of the model are increased.

3) Robotic models both sense and act, which has two significant consequences. First, it means the simulated infant can interact with its environment and "select" its own sensory input. Second, and more importantly, it also means that the model's overt actions – for example, head and eye movements, reaching, pointing, and grasping behaviors, vocalizations, and so on – can be directly compared to those produced by human infants. In other words, the same behavioral measures that are used in an infant study can also be implemented in a robotic model.

4) As we highlight in the discussion, a particularly valuable feature of robotic models is that they not only provide a context for simulating infant experiments, but more generally, they also provide a rich environment for simulating the developmental timescale. Thus, the same robot platform can be used to study both short-term learning and long-term development (i.e., microgenesis and ontogenesis).

To date, only a few studies have used real-world robot platforms to simulate infant-perception experiments. One of the first is [14], which simulated looking-time behavior in an autonomous mobile robot. In this experiment, the robot was presented with possible and impossible versions of an object-occlusion event, and the robot's gaze behavior was analyzed. Like human infants, the robot produced longer looking times toward the impossible event, which it perceived as more novel than the possible event. A related looking-time experiment was simulated by [15], who employed a similar modeling approach and testing paradigm, and reported comparable results. A unique feature of [15], however, is that an upper-torso humanoid robot was used to simulate the young infant.

Given the cost and difficulty of using real robots, a relatively inexpensive and straightforward alternative is to embed the model within a quasi-realistic robot simulator. In particular, the iCub robot simulator (http://eris.liralab.it/wiki/ICub_Software_Installation) provides a comprehensive modeling framework, by not only offering a humanoid robot with the physical characteristics of a 3-year-old child, but also an immersive virtual environment that allows the model-builder to select and precisely control all the properties of the environment (e.g., illumination, surfaces and objects, physical interactions, etc.) [16]. In addition to the features available within the simulator, the *Aquila* (front-end) GUI offers an especially-useful tool: a configurable 2D screen, which can be used to display both static images and animated or digitized events (http://sourceforge.net/projects/aquila/). Thus, as Figure 1 illustrates, the iCub robot can be presented with the same visual displays that are used to test perceptual development in human infants.

III. Perceptual Completion

To provide an example of these ideas in practice, we next present the perceptual-completion task as a specific case study.

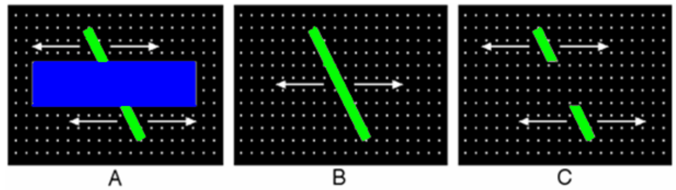

Fig. 2. Displays used to assess perceptual completion in infants: (A) occluded-rod, and (B) complete rod and (c) broken rod test displays.

We first briefly explain how perceptual completion is assessed in young infants, and then describe the eye-movement model that we have designed and tested to simulate the development of this skill.

A. The Perceptual-Completion Task

Perceptual completion is the ability to perceive a partially-occluded object as a coherent, integrated whole. The ability is studied in young infants by presenting events like those illustrated in Figure 2. First, infants view a moving rod that is partially occluded (Figure 2A). After several repetitions of the occluded-rod display, infants gradually lose interest (i.e., their looking-time falls below a predetermined threshold). Infants then view two test events, presented one at a time: a complete-rod display and a broken-rod display (Figures 2B and 2C, respectively).

Given that infants have habituated to the occluded-rod display, longer looking to one of the two test events suggests that event is more novel than or different from the occluded-rod display. Thus, an infant who perceives the occluded rod as a single object should look longer at the broken-rod display. Conversely, an infant who perceives the occluded rod as two disjoint objects should look longer at the complete-rod display. We refer to infants who show each of these looking patterns as *perceivers* and *nonperceivers*, respectively.

The perceptual-completion task has provided a number of major insights into how perceptual completion develops. First, between birth and age 2 months, infants are typically classified as nonperceivers, as they look longer at the complete-rod test display (Figure 2B) [17]. By age 4 months, most infants are classified as perceivers, as they look reliably longer at the broken-rod test display (Figure 2C), which suggests that they perceive the occluded rod as a single, unified object [18].

Second, the scanning pattern produced by a 2- or 3-month-old during the occluded-rod display can be used to predict whether they will be classified as a perceiver or nonperceiver [19]-[20]. In particular, while viewing the occluded-rod display, perceivers generate significantly more *rod scans* (i.e., fixations toward the moving rod segments) than nonperceivers. This finding supports the idea that increased attention toward the moving rod enables the perception of it as a single object.

B. The Eye-Movement Model

In order to better understand these findings, and more generally, to provide a platform for simulating perceptual development across a broad variety of experimental paradigms, we have designed and tested an eye-movement

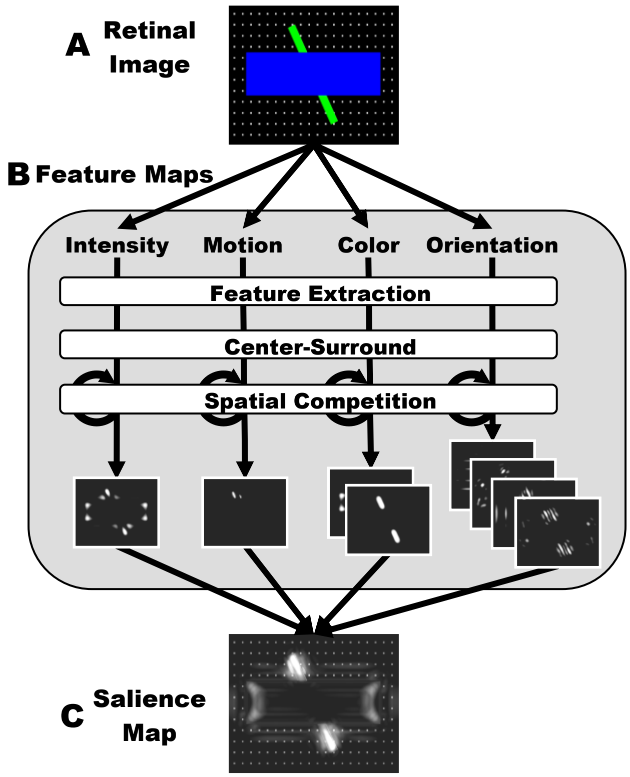

Fig. 3. Schematic diagram of the eye-movement model: (A) An input image is projected onto the retina, (B) the retinal image is projected through four feature channels (intensity, motion, color, and orientation), and (C) feature maps produced across the four feature channels are pooled into a single, unified salience map.

model [3]-[5]. The model is an extension of the salience-map framework proposed by [21]. Processing within the model occurs over four stages, three of which are illustrated in Figure 3 (for additional details, see [4]):

1) *Retinal Image (3A)*. An input image is projected onto the model's simulated retina.

2) *Feature Maps (3B)*. The retinal image projects through four optical filters (i.e., intensity, motion, color, and oriented edges), which decompose the input image into a set of retinotopic feature maps. During this stage, a spatial-competition (i.e., 2D Difference-of-Gaussians) filter is applied to each feature map.

3) *Salience Map (3C)*. The feature maps are pooled into a single salience map.

4) *Eye Movement (not illustrated)*. A highly-active location on the salience map is selected probabilistically, and a "virtual" eye movement to this location is produced.

It should be noted that because the model produces virtual (or covert) eye movements – that is, the field of view remains fixed, while the sequence of target locations is recorded (see [21] for a comparable approach) – it straddles the border between retinal-image and active-vision models (see Sections IIB and IIC).

The eye-movement model has succeeded in replicating infants' performance not only on the perceptual-completion task, but also on a second task that measures visual search [3]-[4]. In particular, we have demonstrated that by hand-tuning a critical parameter in the model, which modulates the amount of spatial competition that occurs within the feature maps (see Figure 3B), the model generates gaze patterns that correspond to the patterns produced by 3-month-old perceivers and nonperceivers, respectively, in the perceptual-completion task.

To help illustrate this finding, Figure 4 presents the proportion of rod scans produced by the model, as a function of the duration of spatial competition. Note that the spatial-competition filter is represented within the model as a discrete, iterative process, which can be applied an arbitrary number of times to each feature map. The red curve in Figure 6 represents the performance of the "retinal-image" version of the model, that is, our prior implementation in which the occluded-rod display is projected directly on to the model's simulated retina (see Figure 3A).

In order to compare the performance of the model with human infants, Figure 4 also includes the proportion of rod scans produced by 3-month-old perceivers and nonperceivers (as reported by [19]): the thin dotted line, at 0.13, is the proportion of rod scans produced by nonperceivers, while the thick dashed line at 0.19 is the proportion produced by perceivers. In the model, rod scans monotonically increase as spatial competition increases, and specifically, the model matches the performance of nonperceivers and perceivers between 3 and 4 iterations of the spatial-competition filter.

These simulation findings highlight the role of preattentive competition among salient locations in early visual processing, and also suggest that increases in spatial competition may help explain the development of perceptual completion in young infants. However, recall that we noted some of the limitations of retinal-image and active-vision models in Section IID, and therefore, the next section presents the simulation findings

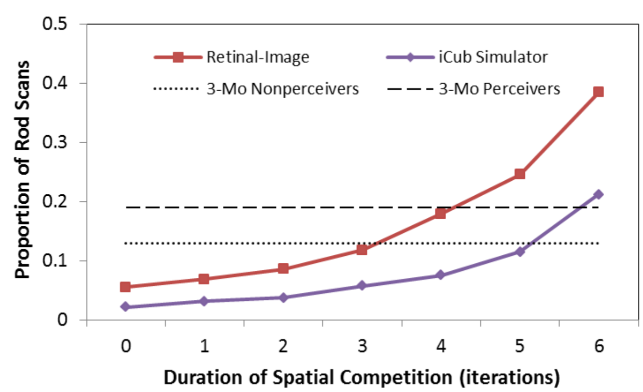

Fig. 4. Proportion of rod scans produced by the model as the duration of spatial competition is increased. The red curve presents the performance of our previous "retinal-image" model, while the purple curve is the performance when the model is embedded in the iCub simulator. The dotted and dashed horizontal lines indicate the performance level of 3-month-olds on the same task.

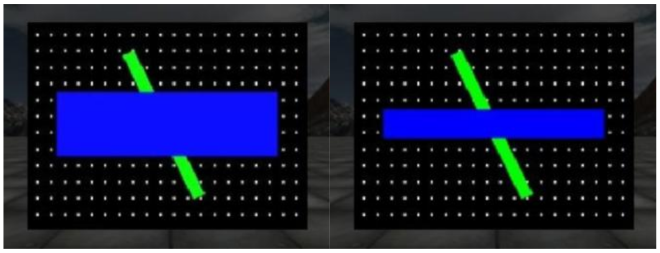

Fig. 5. Screenshots of the canonical occluded-rod display from the iCub's perspective (left image), and the modified short-screen version of the display (right image).

from porting and testing the model in the iCub simulator.

IV. iCub Simulation Results

The right image in Figure 1 presents a third-person view within the iCub simulator of the iCub robot and the projection "screen" used to display the occluded-rod event. As Figure 1 illustrates, we selected the size and placement of the screen so that it was approximately comparable to the spatial arrangement used to test human infants. Figure 5 presents two snapshots of the occluded-rod event, captured from the iCub's left camera (or eye). In the image on the left, iCub views the "standard" or canonical version of the occluded-rod event (Simulation 1), while the image on the right presents a modified version of the event with a shorter screen (Simulation 2).

Because our primary goal was to directly compare the simulation findings from the "retinal-image" and iCub implementations of our model, we chose to hold iCub's torso, head, and eyes fixed during the simulation. Thus, the two versions of the occluded-rod event were projected into the simulator and were recorded (at approximately 25fps) using iCub's left camera. These recordings were then processed offline by the eye-movement model, using the same model architecture and algorithm described in Section III. Rod scans were also defined and measured in the same way as described above.

A. Simulation 1: Standard Screen

The purple curve in Figure 4 presents the proportion of rod scans produced by the model in the iCub simulator, as a function of the duration of spatial competition. There are three important results. First, as in the retinal-image model, increasing spatial competition in the iCub model also results in a monotonic increase in rod scans. Second, however, for each iteration of spatial competition, the model produces fewer rod scans in the iCub simulator than in the retinal-image model. One possible reason for this difference is that because the occluded-rod display does not fill iCub's visual field, the relative salience of locations in the display are somewhat diminished in the simulator.

Third, and perhaps more importantly, the performance of the model in the simulator also matches the level of rod scans produced by 3-month-old perceivers and nonperceivers (at 5 and 6 iterations of the spatial-competition filter, respectively). Thus, the findings from Simulation 1 provide clear support for our model, and demonstrate that it can be successfully ported into the iCub simulator.

B. Simulation 2: Short Screen

Next, we extended our findings by presenting iCub with a modified version of the occluded-rod event. In particular, 2-month-olds are usually classified as nonperceivers when they view the canonical display. However, [22] hypothesized and confirmed that by shortening the occluding screen and making the moving rod more salient, 2-month-olds should shift and respond like perceivers. In a follow-up study, [23] found, as expected, that presentation of the short screen increased the proportion of rod scans produced by 2-month-olds.

We therefore also predicted that presenting the model with the short-screen version of the occluded-rod event (see Figure 5, right image) would increase the proportion of rod scans. The green curve in Figure 6 (left) presents the proportion of rod scans produced by the model (in the simulator) while viewing the short-screen display; for comparison purposes, the purple line re-presents the model's performance with the standard screen. As Figure 6 illustrates, our prediction was partially supported: for low levels of spatial competition (i.e., between 0 and 2 iterations), the proportion of rod scans increased. To help highlight this result, on the right side of Figure 6 we present the proportion of rod scans produced by the model as a ratio of the short-screen versus the standard-screen conditions. The ratio of short-screen to standard-screen

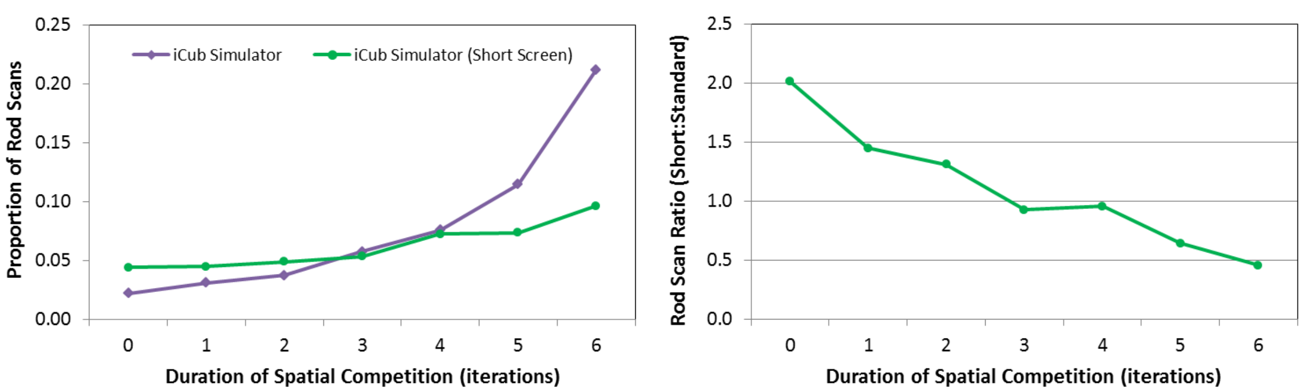

Fig. 6. (Left) Proportion of rod scans produced by the model in the iCub simulator, while viewing the standard-screen and short-screen versions of the occluded-rod display (purple and green curves, respectively. (Right) The ratio of rod scans in the short-screen versus the standard-screen conditions (i.e., short/standard).

rod scans is above 1 between 0-2 iterations, is near parity for 3-4 iterations, and falls below 1 after 4 iterations.

While this pattern was somewhat unexpected, in hindsight it appears to be a straightforward consequence of how the spatial-competition filter is implemented: in particular, the filter implements a "Mexican hat" structure, in which the center of the filter (i.e., convolution kernel) is narrow, steep, and excitatory, while the rest of the filter is broad, shallow, and inhibitory. Thus, because more of the moving rod is exposed in the short-screen display, it also means that with increasing iterations of the spatial competition loop, more of the two rod segments begin to inhibit or compete with each other (and perhaps also themselves, i.e., self-suppression). In behavioral terms, this result may also make sense as it suggests that increasing the salience of the occluded rod may only increase the proportion of rod scans for nonperceivers, but not necessarily for infants who already respond to the standard display as perceivers.

V. Conclusion

Our initial findings with the iCub simulator suggest three important conclusions. First, the eye-movement model we have developed required virtually no modifications as we shifted from the retinal-image/active-vision approach to the more ecologically-valid environment offered by the simulator. Our next step is to use the targets selected from the salience map as input into iCub's gaze-control system, which means that iCub will shift its gaze from one salient location to another. As a result, this will "unlock" a valuable feature of the simulator: by producing head and eye movements, we can also directly measure iCub's looking time (i.e., toward the display versus other locations).

Second, the results of the two simulations not only replicate our previous findings, but also extend them by demonstrating that increasing the salience of the moving rod selectively increases rod scans in the model (i.e., when "endogenous" spatial competition is low). An important limitation of our model, however, is that its performance is hand-tuned. We are currently investigating the use of prediction-learning as a mechanism for tuning the model autonomously [5].

Finally, and more generally, our work only scratches the surface in terms how the iCub simulator can be used as a research tool. In particular, we have illustrated its use for simulating an object-perception experiment. Our long-term goal, however, is to allow the virtual infant to "live and learn" in its simulated environment, so that we can better understand the role of everyday action and experience as a fundamental influence on perceptual and cognitive development.

References

[1] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE Transactions in Autonomous Mental Development*, vol. 1, pp. 12-34, 2009.

[2] J. Weng. J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur , and E. Thelen, "Artificial intelligence: Autonomous mental development by robots and animals," *Science*, vol. 291, pp. 599-600, 2001.

[3] M. Schlesinger, D. Amso, and S. P. Johnson, "The neural basis for visual selective attention in young infants: A computational account," *Adaptive Behavior*, vol. 15, pp. 135-148, 2007.

[4] M. Schlesinger, D. Amso, and S. P. Johnson, "Simulating infants' gaze patterns during the development of perceptual completion," in *Proc. Seventh International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Sweden, 2007, pp. 157-164.

[5] M. Schlesinger, D. Amso, and S. P. Johnson, "Increasing spatial competition enhances visual prediction learning," in *Proc. of the International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob-2011)*, New York, 2011.

[6] D. Mareschal and R. M. French, "A connectionist account of interference effects in early infant memory and categorization," in *Proc. Nineteenth Annual Cognitive Science Society Conference*, Palo Alto, CA, 1997.

[7] Y. Munakata, "Infant perseveration and implications for object permanence theories: A PDP model of the AnotB task," *Developmental Science*, vol. 1, pp. 161-211, 1998.

[8] D. Mareschal, K. Plunkett, and P. Harris, "A computational and neuropsychological account of object-oriented behaviours in infancy," *Developmental Science*, vol. 2, pp. 306-317, 1999.

[9] Y. Munakata, J. L. McClelland, M. H. Johnson, and R. S. Siegler, "Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks," *Psychological Review*, vol. 104, pp. 686-713, 1997.

[10] A. Franz and J. Triesch, "A unified computational model of the development of object unity, object permanence, and occluded object trajectory perception," *Infant Behavior and Development*, vol. 33, pp. 635-653, 2010.

[11] D. Mareschal and S. P. Johnson, "Learning to perceive object unity: A connectionist account," *Developmental Science*, vol. 5, pp. 151-185, 2002.

[12] M. Schlesinger and D. Parisi, "The agent-based approach: A new direction for computational models of development," *Developmental Review*, vol. 21, pp. 121-146, 2001.

[13] M. Schlesinger, "Decomposing infants' object representations: A dual-route processing account," *Connection Science*, vol. 18, pp. 207-216, 2006.

[14] Y. Chen and J. Weng, "Developmental learning: A case study in understanding 'object permanence,'" in *Proc. Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Sweden, 2001, pp. 35-42.

[15] A. Lovett and B. Scassellati, "Using a robot to reexamine looking time experiments," in *Proc. of the Third International Conference on Development and Learning*, San Diego, CA, 2004, pp. 284-291.

[16] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano, "The iCub humanoid robot: An open-systems platform for research in cognitive development," *Neural Networks*, vol. 23, pp. 1125-1134, 2010.

[17] A. Slater, S. P. Johnson, E. Brown, and M. Badenoch, "Newborn infants' perception of partly occluded objects," *Infant Behavior and Development*, vol. 19, pp. 145-148, 1996.

[18] S. P. Johnson, "Development of perceptual completion in infancy," *Psychological Science*, vol. 15, pp. 769-775, 2004.

[19] D. Amso, and S.P. Johnson, "Learning by selection: Visual search and object perception in young infants," *Developmental Psychology*, vol. 42, pp. 1236-1245, 2006.

[20] S. P. Johnson, J. A. Slemmer, and D. Amso, "Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds," *Infancy*, vol. 6, pp. 185-201, 2004.

[21] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489-1506, 2000.

[22] S. P. Johnson, "Development of perceptual completion in infancy," *Psychological Science*, vol. 15, pp. 769-775, 2004.

[23] S. P. Johnson, J. Davidow, C. Hall-Haro, and M. C. Frank, "Development of perceptual completion originates in information acquisition," *Developmental Psychology*, vol. 44, pp 1214-1224, 2008.