MORPHOLOGICAL SHAPE ANALYSIS OF CHILDREN'S ACTIVELY GENERATED VIEWPOINTS

Alfredo F. Pereira[1], Isabel C. Lisboa[1], Emanuel Sousa[2], Jorge A. Santos[1,2,3], Karin James[4] & Linda B. Smith[4]

[1]*School of Psychology, University of Minho (Portugal)*

[2]*Centro Algoritmi, School of Engineering, University of Minho (Portugal)*

[3]*Centro de Computação Gráfica (Portugal)*

[4]*Psychological and Brain Sciences, Indiana University Bloomington (USA)*

Abstract

Our visual system develops in a world of three-dimensional objects, where the dynamic visual experience of objects is often under the child's active control. Previous experimental work by some of us was the first to analyze in detail what kind of object views are selected in children's unconstrained manipulation of physical objects (Pereira, James, Jones, & Smith, 2010). The key finding was the distribution of dwell time across all possible viewpoints: preferred viewpoints are around on-axis views, where the principal axis of elongation is either perpendicular or parallel to the line of sight and flat surfaces are presented approximately perpendicular to the line of sight – so called planar views.

Here we present a morphological shape analysis of the object views selected by a cross-sectional sample of children that manipulated objects in the Pereira et al. (2010) study (N = 54, 12-36 months). We followed an image-based (pixel-based) approach; we coded 3D object orientation and created computer-generated images of the object; finally, we extracted two shape measures from the generated images. In particular we computed surface area and aspect ratio. For the main analysis we investigated if the shape measures differ between planar and non-planar views. In addition, and for a smaller subset of 30-36 month olds (N = 8) we computed the number of pixels that do not overlap between consecutive frames (a measure of image variability).

There were two key results: (1) on average, planar views have lower surface area (only in the oldest age groups), and aspect ratio increased with developmental age; and (2) on average, consecutive frames of planar views have a lower proportion of pixels that do not overlap, compared with non planar views. We discuss these results in the context of their contribution to our understanding of the development of visual object perception.

Keywords: *Object recognition, Perceptual development, Perception and action, Visual development, Active object inspection.*

1. Introduction

Children influence what they see as they move their eyes, their head, their body, and as they manually explore and act on objects. Manipulating an object – a privileged form of visual stimulation – is an opportunity to perceive changes to the visual image contingent on self-generated movements. Underscoring the importance of active visual learning, object pose under active control is highly biased: adults and children commonly select or prefer certain views and do so far from chance selection.

Recent work has analysed children's self-generated object views as they manipulate objects (James, K.H., Jones, Swain, Pereira, & Smith, in press; Pereira et al., 2010). The first study, Pereira et al. (2010), was a cross-sectional study of 12 to 36 month old children that examined two structural and important aspects of visual object learning: (1) which views children prefer over others, and (2) how often objects were in an upright position. Participants explored objects in a free-play task and 3D orientation was measured. Analysis of dwell time across all possible object views revealed significant developmental changes in the structure of children's active viewing; the first was an increasing bias for planar views – an object pose where the main axis of elongation is, approximately, either orthogonal or parallel to the line of sight and flat surfaces are also presented approximately orthogonal to the line of sight (Harman, Humphrey, & Goodale, 1999; Perrett & Harries, 1988). The second structural property concerns how the object is held in respect to its upright position – 30 to 36 month olds are biased towards views that present the object in an upright position.

The planar bias in particular emerges developmentally early (measurable in tge 12 to 18 month old age group) and increases in strength until 36 months (Pereira et al., 2010). This effect is well-documented in the adult literature and has been replicated across stimulus and tasks (see Pereira et al. (2010) for a discussion) and critically, one study suggests it promotes more efficient visual object learning (James, K. H., Humphrey, & Goodale, 2001). These findings show that active perceivers, even very young children, constrain the problem of visual object learning with specific viewing biases but these also change considerably early in development. The functional role of the planar bias is still poorly understood and we known little about the properties that distinguish preferred viewpoints from other object views.

We present here the findings of a pixel-based morphological shape analysis of the Pereira et al. (2010) study's data. We examined if planar views differ in terms of surface area and aspect ratio from non-planar views, and if the exploration around planar views is more stable by measuring the difference between consecutive frames.

2. Study 1: Image-based analysis of actively selected viewpoints

2.1 Participants

Children aged 12 to 36 months participated in the Pereira et al. (2010) study. Participants were divided in four age groups: 12 to 18 months (N = 14); 18 to 24 months (N = 14); 24 to 30 months (N = 14); and 30 to 36 months (N = 12).

2.2. Procedure

For the Pereira et al. (2010) study, children came to the laboratory for a free object play task. Participants manipulated objects, one at a time, and without a supporting surface, while wearing a lightweight mini head-mounted camera, low on the forehead and calibrated so the object would be in the center of the image. During this developmental period, and particularly in a context of object manipulation, eye and head directions are typically aligned (Yoshida & Smith, 2008), and thus the head-mounted camera method can be used to capture the child's self-generated object views. The stimulus set was composed of objects from eight common object-based categories in two levels of shape detail (16 objects in total); subjects saw one object per category.

Object orientation was coded at 1Hz using a custom-made software application that allowed a human coder to observe a video frame from the mini camera and manipulate a 3D model of the object until it matched the image – the model's 3D orientation was recorded as the estimate of the child's view. An object view was coded as planar view if the angle between the line of sight and the normal vector of the object's bounding box was zero (with a tolerance of ±11.25 degrees, see Figure 1 left panel). Complete details on stimuli, the experimental procedure (including the calibration of the head mounted camera) and the 3D orientation coding can be found in Pereira et al. (2010).

The 3D orientation data from the previous cross-sectional study was the basis for the present study. Using the 3D models and the orientation data, the manipulations were reconstructed by rendering one image per coded head-mounted camera video frame: 12 to 18 months (N = 1891 frames); 18 to 24 months (N = 1680 frames); 24 to 30 months (N = 1803 frames); and 30 to 36 months (N = 2421 frames). The images were further processed using Matlab in two steps: first, they were binarized – all object pixels were considered on and all background pixels were considered off; and second, the function *regionprops* was used to compute the object blob image properties (see Figure 1 right panel). We examined two measures, total surface area (provided by *regionprops*) and aspect ratio. The later was expressed as the aspect ratio of the ellipse with the same normalized second central moments of the image blob (an output of *regionprops*), calculated by dividing the ellipse's major axis length by the ellipse's minor axis length – an aspect ratio of 1 is a circle; the larger the aspect ratio, the more elongated the ellipse. The total surface area was normalized using the maximum object-specific area present in the dataset, and the aspect ratio was range normalized (subtracting the minimum value and dividing by the object-specific aspect ratio range).

Figure 1. (left panel) Planar view definition: a view was considered planar if the normal vector of object's bounding box made an angle of 0º±11.25º with the line of sight – in the example, LoS_1 corresponds to a planar view and LoS_2 to a non-planar view. (right panel) An example of a sequence of object views generated by a child and reconstructed using computer renderings of the objects; the bottom row shows the rendered frame in binary format.

In addition, we also recoded a sub-set of the data at a higher sampling rate (10Hz) in order to be able to compute a measure of image stability. Only N = 8 participants, randomly selected from the oldest age group, were coded, for a total of 14446 frames. The number of pixels that don't overlap between consecutive frames was calculated and normalized by the maximum surface area in pixels per object.

2.3. Results

Children were free to manipulate the objects for as long as they wished and older children spent more time inspecting objects than younger children. The mean holding times per age group were, in increasing age period: 15.4, 22.0, 27.5, and 34.7 seconds respectively.

The first analysis consisted of fitting the normalized surface area and the range normalized aspect ratio data with a 4 *AgeGroup* x 2 *ViewType*(planar/non-planar) multivariate general linear model, with both measures entered as dependent variables, and *AgeGroup* and *ViewType* entered as between-subjects factors. Post-hoc analysis was conducted by examining all 2-way pairwise comparisons using Sidak correction. Figure 2 shows the observed sample cell means for this design.

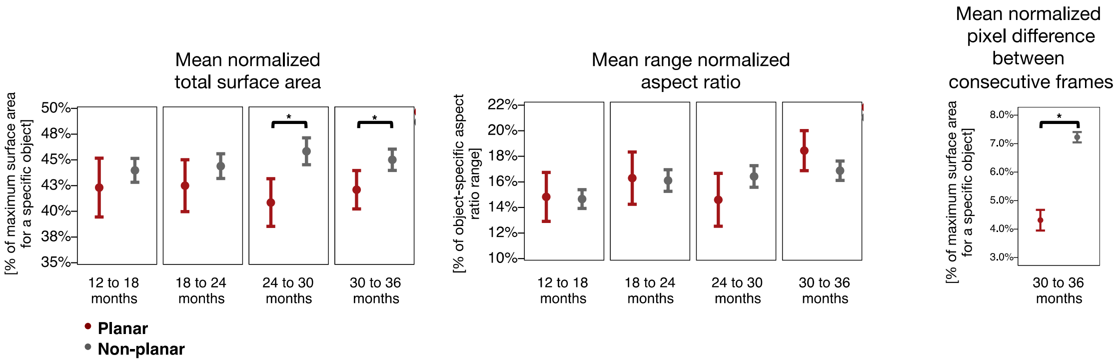
This analysis yielded a main effect of *ViewType* for normalized surface area, $F(1, 7787) = 15.77$, $p < .001$, and a main effect of *AgeGroup* for range normalized aspect ratio, $F(3, 7787) = 7.21$, $p < .001$; all other effects were non-significant ($p > .05$). On average, actively selected object viewpoints close to a planar view corresponded to a lower object surface than non-planar views (Figure 2 left panel). Post-hoc comparisons revealed this difference was only significant in the two older age groups. For the aspect ratio measure there were no average differences between planar and non-planar views across all age groups; however, there was a general increase over age for this measure – older children chose more elongated views (Figure 2 middle panel). Only the difference between the 30-36 months and the 12-18 months or the 24-30 months age groups was significant, according to the post-hoc comparisons.

The normalized pixel difference measure revealed that – at least for the oldest developmental age period where we computed this measure – planar views correspond to moments where the pixel difference between frames is considerably

ISBN: 978-989-97866-6-0 © 2014

lower than the rest of the object manipulation, $t(3032.2) = 14.37$, $p < .001$, for unequal variances (Figure 2 right panel) – i.e. exploration around planar views corresponds to moments where the image was more stable.

Figure 2. Study 1's results. Sample average values, per age group, for normalized surface area (left panel) and range normalized aspect ratio (middle panel). Solely for the oldest age group, 30 to 36 months, the right panel shows the normalized pixel difference between frames. Data is shown for planar and non-planar views; error bars mark 2 ± SEM; and the asterisk marks a significant comparison ($p < .05$)



3. Discussion

Children actively co-construct the natural statistics of their own visual experience of objects during a long period of object exploration – and thus shape the development of the visual system in important ways (Smith & Pereira, 2009). The presence of biases in active visual learning, and in particular the planar bias discussed here (a bias in the sampling of object viewpoints), suggests the presence of visual learning mechanisms at work and may contribute to understanding the molar features of shape used by the visual system. Consistent with this hypothesis, our previous study found the planar bias to be developmentally early but also to increase in strength considerably (Pereira et al., 2010).

Here we investigated if the planar viewpoints, selected by children during object manipulation, differ in terms of two basic static shape properties, from non-planar object views. We measured total surface area and aspect ratio (a measure of elongation); the planar bias is about orienting surfaces so they are approximately orthogonal to the line of sight, and thus these two measures were relevant in testing if simple differences in the characteristics of the surfaces in view would explain the planar bias. In addition, and because one hypothesis for the planar bias' functional role is the exploration of viewpoint instabilities (James, K. H. et al., 2001; Pereira et al., 2010) – rotations around planar viewpoints typically generate larger changes in the available object information than around non-planar viewpoints – we also investigated a measure of image stability: pixel difference between frames (this measure only for a subset of the oldest age group).

The results, using a stimulus set designed to sample widely from common object shapes, show that planar views do differ in surface area but they correspond to object views with less visible surface. This difference is, however, not present early on (only children older than 24 months show this difference) which stands in contrast with the planar bias that can be observed in the earliest age group tested so far (12 to 18 months). A simple explanation of the visual bias based on expansion of surfaces seems unlikely. One possibility, nevertheless, is that younger children – with a weaker planar view bias (there is a 1.5 fold increase in the proportion of planar views between the youngest and oldest age group in the Pereira et al. (2010) study) – orient objects less precisely and thus generate planar views that are more similar to non-planar views. The surface area finding is also relevant because it supports another important contrast, one linking visual object learning and visual object recognition. Canonical views in recognition are typically not planar (the main axis of elongation is oblique) and show multiple sides; active perceivers, nevertheless, prefer views that are of intermediate levels of complexity (they show only one side of the object).

Regarding aspect ratio, although planar viewpoints require the alignment of the main axis of elongation (and therefore one might predict more elongated views), the elongation is not reliably present in planar views, although there is an increase in aspect ratio with developmental age. Additional studies are needed, with stimuli specifically designed to test this property.

The finding on dynamic viewing around planar views, regarding the oldest age group, is an important novel result. The lower pixel difference between frames around planar views means they correspond to more stable periods of the object manipulation, in opposition to initial conjectures associating the planar bias with higher instabilities (caused by movements of exploring planar viewpoints). This result suggests that sampling planar views corresponds to moments of focused attention to a particular view, where perhaps learning of a static view is occurring. However, it is still possible that participants explored more finely the planar views (and thus generated purposeful movements around those views) and this produced the observed difference. Further work is required to elucidate the detailed nature of the motor movements involved. Critically, the present results and approach leave open the question of how view transitions are integrated over time.

References

Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9*(22), 1315-1318.

James, K. H., Humphrey, G. K., & Goodale, M. A. (2001). Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology, 55*(2), 111-120.

James, K. H., Jones, S. S., Swain, S., Pereira, A. F., & Smith, L. B. (in press). Some views are better than others: Evidence for a visual bias in object views self-generated by toddlers. *Developmental Science*.

Pereira, A. F., James, K. H., Jones, S. S., & Smith, L. B. (2010). Early biases and developmental changes in self-generated object views. *Journal of Vision, 10*(11).

Perrett, D. I., & Harries, M. H. (1988). Characteristic views and the visual inspection of simple faceted and smooth objects: "Tetrahedra and potatoes.". *Perception, 17*(6), 703-720.

Smith, L. B., & Pereira, A. F. (2009). Shape, action, symbolic play, and words: Overlapping loops of cause and consequence in developmental process. In S. P. Johnson (Ed.), *A neo-constructivist approach to early development*: New York: Oxford University Press.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy. Vol, 13*(3), 229-248.