Human-Level AI and Comptational Cognitive Neuroscience *Yezhou Yang*

Writing Assignment 3: SPELKE

1 Who is Elizabeth S. Spelke

Dr. Elizabeth S. Spelke is a professor of psychology at the Harvard University Laboratory for Developmental Studies. The main focus of her study, as I quote from the interview "Insights from the Youngest Minds" [1], is "study babies".

Actually despite of conducting experiments on newly born babies, she is trying to identify the bedrock categories of human knowledge, like what is number, space, agency, and how does knowledge in each category develop from its minimal state. Questions about human cognition and the organization of the human mind fascinate her but adult mind is too complicated to study, "too stuffed full of facts" to make sense of it. In her view, the best way to determine what humans are born knowing, is to go straight to the source, and consult the recently born.

The tool of Dr. Spelke and her colleagues mainly rely on is the use of the infant gaze as a key to the infant mind. Study human gaze has long been a popular way to decode what is happening in human mind. However, to adults, it is not always straightforward to analyse their gazes to what they are thinking. For example, a lot researchers working on visual attention, and by analysing human gaze distribution, researchers can make some claim like "human being tend to attend on visual input with high contrast than low contrast ones". But apparently since adult's mind is "too stuffed full of facts", the life experience they have is biasing the experimental results dramatically. Imagine you put an image with a handsome man and a beautiful woman walking along the way, disregard of whether they are wearing high or low contrast outerwear, a male tester attend more on the female in the pic, and vice versa to the female tester.

That's why Dr. Spelke and her team study infant gaze, and they are pioneers of it. The way to conduct infant gaze based experiment is by identifying the inherent expectations of babies as young as a week or two by measuring how long they stare at a scene in which those presumptions are upended or unmet. We introduce one experiment done by this method in the next section.

2 Infant Object Permanence

One of the most famous experiments they conducted in [2] is a typical example of using infant gaze to study developmental object permanence capability.

In the experiments, infants around five-month-old were shown two test events: a possible event and an impossible event. In the possible event, the screen stopped when it reached the occluded box; in the impossible event, the screen move through the space occupied by the box. The infants turned out to look reliably longer at the impossible than at the possible event, which suggests that the infants (1) understood that the box continues to exist, in its same location after it was occluded by the screen, and (2) expected the screen to stop against the occluded box and were surprised, or puzzled, when it failed to do so.

The beauty of this experiment in my view lies in the elegant design, where the researchers successfully utilize infants gaze to justify some presumptions.

3 My Thoughts

As a computer science student, and my research interest lies in Computer Vision and Robotics, it is always inspiring by reading those psychology experiments, and I really enjoyed reading them. In my view, nowadays people in Computer Vision community is struggling to identify themselves as engineers or scientists. It puzzles me for a long time also. For example, when talking about object recognition, from a scientific research point of view, the questions to be asked should be like what defines an object? (In [2], they claim that for adults, an object is an entity that exists continuously in time and space: it cannot exist at two separate points in time without having existed during the interval between them, and it cannot appear at two separate pints in space without having travelled from one point to the other) what makes an object conceptually different from another object? What are the features human beings relying on to discriminate between conceptually different objects?

However, the goal of object recognition in Computer Vision is that given an image or several images of an object, develop a system to point out what kind of object it is. Thousands of methods are published in the last several decades, with all kind of different features proposed and method used. The most famous and well recognized as an effective method is a machine learning method (SVM) with histogram of gradient (HoG) feature. It provides reliable recognition performance on benchmark datasets.

Here comes the question, why HoG is such a discriminative feature for objects? Is it because human being are naturally better at sensing gradient than actual colors, or objects are generally regarded as belonging to different classes when they have different gradient distribution? I think we can design some experiments using infant gaze to test. For example, prepare two sets of images, one set are two apples with different colors (say one green, one red), and another set are two apples with different texture (say one unicolor, another striped). The objects in the first set has similar HoG representation, while those in the second set has dramatically different HoG. We show the infant with one of the image in each set at the beginning, then use a box to occlude them. Since from [2], we know that even child at five-month-old knows that the object continue to exist if occluded. Now we secretly change the object occluded to the other image in each set. Now by comparing how surprise or how long the infants gaze at the later object, we can verify that how different the infant think those objects are, and whether they think the object is dramatically changed, which violates their assumption, and tends to put them into two different conceptual classes. If infants gaze reliably longer at the second set, then it suggests at the beginning, we tends to differentiate objects into different classes by their gradient distribution (at least more than just using color itself). If that is the case, it in some way supports why HoG is successful in object recognition tasks.

In other words, the question is, whether perceptually different objects are also conceptually different, and which way of perception differentiate objects conceptually. Taking Zibra as an example, we call horse like animal with striped texture as Zibra, but we seldom make up a new word for horses with different uni-colors, even the color is extremely eccentric, like snow-white. Is it also a cue to show that human beings by their nature, tends to classify objects into different conceptual classes by gradient distribution? What if there exists an animal like that Photoshoped one in Fig. 1? We may already have a fancy new name for it, maybe Zirog?

The frustrating thing is that researchers would rather spend time on doing slight change of the HoG, or using different fancy learning methods to improve even one percent of the performance. Very few people cares why those features are good by its nature, or why a specific learning method performs better.

4 References

[1] “Insights From the Youngest Minds”, Natalie Angier, New York Times, April 30, 1012;

[2] “Object permanence in five-month-old infants”, R. Baillargeon, E. Spelke, S. Wasserman, Cognition, 20 (1985) 191-208

Figure 1: Zibra? Zirog?