CIS 581
Project 3B Proposal
Ariana Familiar (Group #24)

Abstract

Humans have the remarkable ability to remember thousands of objects they encounter and generalize over past experiences to recognize new objects based on defining information. For centuries, philosophers and psychologists have studied how human cognition, and specifically what learning mechanisms, allow for rapid and automatic object recognition. That being said, most studies have been based on small sets of object stimuli (e.g. 20), and examining behavior using real-world objects is confounded by preexisting knowledge of the objects. Some studies have attempted to use novel objects, however such stimuli are contrived and often very simple in visual properties while real objects are much more complex. In the present project, I aim to create novel objects that are based on real-world objects, which can be used to systematically examine how and under what conditions people learn about visual objects. As a Psychology PhD student, I will be able to use such objects as stimuli in future experiments.

My main goal is to develop novel objects that retain visual information about real-world objects and can be used to test and compare recognition abilities in humans. Importantly, these objects will be generated using shape and color information derived from images of real objects. In other words, naturalistic geometric and texture information will be used to produce a set of novel objects.

Related Work

Prior research has established the behavioral relevance and importance of shape as a defining feature of many objects. Shape information is immediately available through visual input and is informative of the functional use of an object and how we may physically interact with it. There is even a specific brain region that contains neurons that are selective for shape information[1]. Thus, the primary feature that will be used to generate objects is shape information derived from the shapes of real-world objects.

Research labs have made publicly available large sets of object images for academic purposes. In the current project, I plan to use a set of images of unique objects from a former study[2] (available at: **https://bradylab.ucsd.edu/stimuli.html)**. I will attempt to use all 2400 images in the set, however depending on how computationally expensive this proves to be I may use a subset of the images.

Although there is a rich literature of research on human object recognition, there has been no former work, to my knowledge, that has attempted to create novel object stimuli based on the statistics of real-world objects.

Proposed Methodology

First, I will convolve each image with a Gaussian blur kernel to remove small details while preserving global shape information. Next, I will use a Canny edge detector to extract the overall shape of each object. Then, I will use morphing and texture synthesis to iteratively combine the shapes and textures of different pairs of objects and generate new objects, as described below.

I will randomly select a pair of objects and warp the shapes of the objects (the outputs of the edge detector) using Delaunay triangulation and thin-plate spline methods; corresponding points will be automatically placed at equal spacing around the border of the shapes. This will produce an average shape between the objects. I then will sample a smaller region within the original object image, and cross-dissolve within samples to create an average texture. I will repeat this process with the averaged shape and texture and a randomly selected new object image for many iterations. After the last iteration, I will use texture synthesis to expand the resulting texture sample to fill the entire shape.

I will repeat this process 50 times, in order to generate a set of 50 novel objects. There will be a constraint that an image may never be processed more than once.

Expected Results

I believe the shapes will appear novel but will retain characteristics of the real-world objects. It will be interesting to test different amounts of morphing (i.e. iterations) and examine how this influences the output of shape and texture combinations. There may be a tradeoff between the appearance of the novelty of the object and the distinctiveness between the objects in the set. For instance, with a lower level of morphing iterations, the stimuli will likely appear more distinct from one another, but may look similar to a real-world object (e.g. may appear like a broom). However, the texture cross-dissolving may prove to eliminate the likeness of a novel object to a real-world object (e.g. its shape may appear broom-like but it may not seem like a broom because of its texture). With more iterations, the novel stimuli may appear more similar to one another, as more objects, and their unique variance in shape and texture, are being averaged over.

Expected Applications / Future Work

This work would promote the use of naturalistic stimuli instead of contrived, simplified stimuli for studies of visual cognition in humans. It would not only establish methods for generating novel stimuli based on real-world objects statistics, but would also allow for systematic testing of what physical properties humans use for object recognition. Furthermore, the same methods could be used to generate novel items based on different object categories (e.g. manmade vs. natural objects), and such stimuli could be employed to examine the physical properties that define real-world categories. Future versions of this pipeline could incorporate 3D object information to generate even more realistic novel stimuli.

This work may also be applied to comparisons of human and computer vision systems, particularly for cases in which human vision continues to supersede computer vision. For instance, one could simply add noise to the novel objects in order to simulate poor viewing conditions (e.g. add blur or occluders) and examine how humans accomplish recognition in such settings. This would better establish how humans succeed in such tasks, and could be used to build better computer vision models.

1. Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*(5534), 1506-1509.
2. Brady, T. F., Konkle, T., Alvarez, G. A. and Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA*, 105(38), 14325-14329.