Introduction: Object identification presents a daunting computational challenge for the brain. Beginning with over a hundred million photoreceptors, the visual system must successively separate signal from noise, extract low-level features across a variety of angles, shadings, and other contextual attributes that may affect appearance, and integrate such features into stable higher-level representations. Prior studies have established a caudo-rostral spatial organization of the ventral stream with increasingly abstracted representations of visual information from primary visual cortex (V1) toward inferotemporal (IT) cortex. This pattern includes more rostral populations displaying orientation and context-invariant receptive fields tuned for object identity.[1] Various frameworks have been developed to describe the specific computations entailed; one proposal argues that invariant object representations should be viewed as manifold learning, that is, visual areas form stable representations of objects by learning a high-dimensional manifold (surfaces with a locally Euclidean structure e.g. a sphere) formed by the object in the retinal feature space and successively flattening the structure.[2]

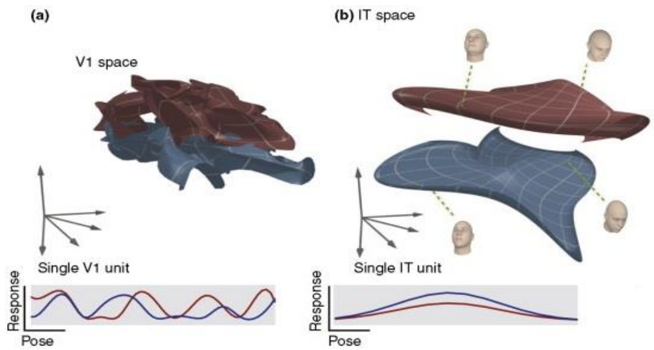

Figure 1: Object manifolds carved out by population responses to variations of the same object.[2] Successive transformations flatten and simplify the manifold as visual information moves along the ventral stream, enabling an easy read-out of identity at IT cortex.

To understand the transformations at work in the ventral stream, I propose using a topological measure of complexity called homology.

Background: In neuroscience, where experimental noise is plentiful and stochastic processes abundant, we often look for stable features of data. In particular, one can analyze the *homology* of a dataset, the number of k-dimensional holes along it. In contrast to geometric properties sensitive to warping of a surface (e.g. curvature, length), topological properties like homology are far more robust as they are preserved under continuous deformations.

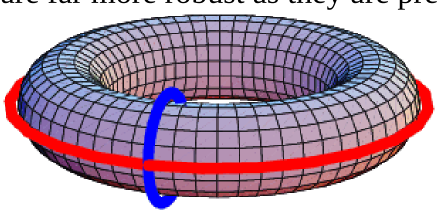

Figure 2: The homology of a torus includes two-dimensional holes along its longitude and latitude. Stretching and bending of the surface does not affect these properties, making them especially robust markers for identifying the torus.

Homology is hence resistant to statistical noise in real-world applications. Examining the homology of a space, we also measure its deviation from linearity (e.g. the torus in Fig. 2 has holes that prevent it from being a plane) and hence its complexity. To minimize complexity, we flatten a surface by manipulating its geometry by continuously bending and stretching appropriately. However, **continuous operations *cannot* remove homology**, presenting a challenge to simplification of object manifolds along the ventral stream. Resolving this problem relies on *non-continuous* cutting and pasting (imagine cutting along the cycles in Fig. 2 to turn the torus into a rectangle) to remove homology.

Prior work concerning increasing linear classifier performance along the ventral stream, as well as computational modeling of V1 and IT population responses, suggest that the transition from a dense to a sparse code along the ventral stream is accomplished by successively flattening and disentangling intertwined object manifolds (Fig. 1, a highly curved object manifold in V1 becomes reduced to a far more linear population response as one moves rostrally toward the temporal lobe).[2] However, it is unknown if this complexity reduction is mediated by changes in manifold geometry or in homology as well. **As these operations are supported by separate classes of functions, their biological implementations might reflect different micro-circuitry**. Several biologically plausible mechanisms have been proposed that could enable continuous warping of object manifolds into flattened surfaces.[2] However, operations that reduce homology are non-continuous in nature and hence evidence of homology reduction would suggest the existence of other algorithms at work. **To probe which operation(s) take place, I propose an analysis of the topological complexity of object manifolds carved out by population responses along the ventral stream.**

Experiment: As object recognition is a rapid process (often <300 ms) and high spatial resolution is necessary for delineating responses of nearby populations, I would use human intracranial recordings in a repeated-measures task, taking advantage my experience with recording paradigms and visual perception.[3] Participants will view a series of images corresponding to the same object category with parametrically varying orientation, shading, and context as well as null trials presenting a separate object. Following each image, participants would categorize the object. I will then analyze response patterns through local field potentials (LFP) in ventral stream subregions. For a given subregion, I will average activity across a post-stimulus window and concatenate across electrodes, generating a summary population response that characterizes a single point on the object manifold (LFP equivalent of Fig. 1 manifolds). By recording population responses to variations of the same object category, we sample along the manifold. I will then deploy an algorithm (persistent homology; from personal statement) to reconstruct the underlying manifold from the point approximations and analyze its homology.[4] Performing this for all regions, spatial trends in homology reduction can then be visualized and tested against the null hypothesis that topological complexity is constant across the ventral stream.

Broader Impacts:

Scientific: The experiment would directly test a theoretical model of ventral stream processing and shed light on the range of possible algorithms implemented. A regional analysis would enable functional mapping of ventral stream regions: I could identify regions located at highly compressive stages of manifold processing, allowing for assignment of putative roles or relative importance in manifold flattening. The resulting framework could then be applied to analyze functional convergence or divergence of visual regions across species, in turn helping to infer evolutionary relationships and generating rich datasets for future work. I would also examine how error trials might relate to a breakdown in either topological or geometric flattening.

Societal: The results would have implications for biologically-inspired machine vision research by characterizing the types of computations at work in biological vision, thereby restricting the space of possible algorithms to be investigated. Machine vision is rapidly impacting fields ranging from medicine to finance, enabling automated diagnosis of diseases from medical imaging data and digit recognition for check deposits.[5] My proposal therefore has the broader impact of accelerating this progress. Additionally, the results would present an opportunity to educate the public about object recognition, an everyday function.

---

[1] Reddy, L, & N. Kanwisher, (2006). *Current opinion in neurobiology*. [2] DiCarlo, J., & D. D. Cox, (2007). *Trends in cognitive sciences* [3] Helfrich, R. F., Huang, M., Wilson, G., & Knight, R. T, (2017). *PNAS.* [4] Varshney, K., and K. N. Ramamurthy, (2015). *IEEE*. [5] Rawat, W., & Wang, Z, (2017). *Neural Computation*.