

Representational geometry of hierarchical category structures in the monkey inferotemporal cortex

Abstract: Psychological theories suggest that humans categorize objects hierarchically based on abstraction of common and shared attributes (1). Prior studies have also found evidence for hierarchical representation of visual categories through clustering and representational similarity analyses of neural activity in the monkey and human ventral visual pathway (2, 3). However, the format and organization of these hierarchical representations remain unknown. We address this question by extending the manifold capacity theory, a recent advance that builds on machine learning and statistical physics to connect the geometry of neural representations with their ability to support classification. We apply our analytical approach to the responses of monkey inferotemporal neurons to a large stimulus set selected from a hierarchy of object categories. A neural manifold is defined as the aggregated neural responses to each category. Quantifying changes in manifold geometry across different levels of the hierarchy, we discover a gradual reduction of distances between manifold centers from the lower levels of the hierarchy to the higher levels, accompanied by minimal changes in the width of manifolds. The combination results in reduced separability of superordinate categories (reduced capacity) compared to subordinate and entry-level categories, matching behavioral studies. We further probe the representational geometry of manifolds by manipulating the manifold shape and orientation, and exploring the impact of different manipulations on the manifold capacity. Our results suggest manifold orientation and overlap attribute to the trends we observe, and hierarchical representation of categories stems largely from judicious spacing of their manifolds.

Additional Details

We analyzed responses of inferior temporal (IT) cortex neurons to visual stimuli sampled from 13 categories within an intuitive hierarchical structure (Fig.1A). In this tree, the categories at the lowest level (subordinate level) joined to form four middle-level categories (entry level): faces, bodies, hands, and objects. The entry level categories grouped to form two higher level categories (superordinate level): animate and inanimate. Two monkeys passively viewed rapid serial visual presentations of the stimuli during the recording sessions. We measured the manifold capacity of categories and their accompanying geometries at different levels of the hierarchy. To test if the intuitive hierarchy is present in the dataset, we repeated the analyses with surrogate datasets created by shuffling the category labels at each level of the hierarchy. Our results indicate hierarchical ordering of neural representations formed by gradual changes of manifold properties across the hierarchy. There is distinctly greater capacity in the unshuffled set of manifolds than in the shuffled set across all three levels of the hierarchy (Fig.1B). While the decrease in capacity from the subordinate categories to the superordinate categories is present for both shuffled and unshuffled manifolds, the change in capacity is distinctly greater for unshuffled manifolds.

Higher capacity at the lowest level of hierarchy could stem from smaller manifold widths or larger center norms – the distance of the manifold centers from a global mean. We show that the manifold size has little change across the levels of the hierarchy, whereas center norms substantially decreased from the subordinate to superordinate levels (Fig.1B). These results indicate that the change in capacity is primarily driven by reduced distance of upper level manifolds instead of increased manifold size.

We examined other possible hierarchy structures formed from the 13 subordinate manifolds. The intuitive hierarchy of Fig.1A has one of the highest capacities among all possible hierarchies from the 13 lowest level categories. This higher manifold capacity did not arise from the superordinate categories, which showed similar capacity for all possible ways of forming superordinate categories through arbitrary combinations of the 4 entry-level categories (Fig.1C). Rather, the increased capacity stemmed from how subordinate categories combined to form the entry-level categories. Put differently, the brain seemingly maximizes the separability of the entry-level manifolds, perhaps due to their ethological significance.

The minimal variation in the size of the manifolds across the hierarchy provided an important clue about the organization of manifolds in state space. For example, the ‘face’ manifold is no bigger than any of its children manifolds, e.g., ‘monkey face’ or ‘human face’. How could the combination of lower level

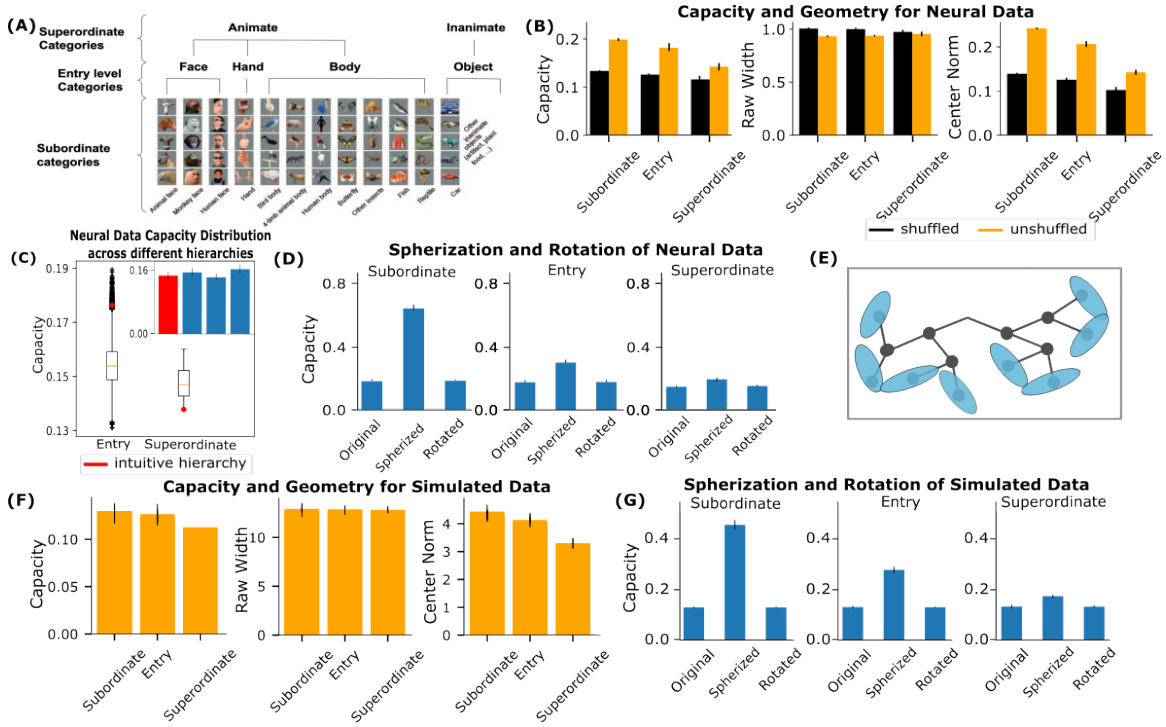


Figure 1: **(A)** The human intuitive hierarchy organization for our 13 manifolds. **(B)** Capacity, mean raw width, and mean center norm for manifolds at each level of abstraction, compared against manifolds with shuffled labels. Center norm describes a manifold’s distance from the global mean. Raw width summarizes the size of the manifold using points on its convex hull that is derived from capacity theory. **(C)** The distribution of capacity for different choices of hierarchies at the entry level and superordinate level. Results for superordinate level is shown in details in the inset plot. **(D)** Change in capacity after spherizing or rotating manifolds at each level of abstraction. **(E)** Our simulated model constructs Gaussian manifolds with an imposed hierarchy. This simulated model replicates the same trends for capacity **(F)** and spherization results **(G)** as the neural data.

manifolds not increase the size of the higher level manifolds? One possibility is that manifolds might be stacked on top of each other along a dimension with low variance. To test this hypothesis, we spherized, or whitened, each manifold while preserving the manifold center position and volume. If manifolds are stacked, spherization would reduce their separability and therefore decrease manifold capacity. To intuit why that would happen, consider a stack of papers. The papers are perfectly separable in the stack, but crumbling them into balls without changing their relative distances would lead to considerable overlap, reducing their separability. Contrary to this prediction, we found that spherization led to increased manifold capacity (Fig.1D). Additionally, we found that randomly rotating the manifolds in place did not alter capacity, providing further evidence against the stacked hypothesis.

Our results give evidence against the stacked hypothesis and support an alternative hypothesis in which manifolds are randomly orientated and overlap at the extremes. Spherization of such manifolds reduces their overlap and improves manifold capacity. Random rotations would not alter capacity as they lead to similar overlaps across manifolds. And finally, under certain conditions, combining the subordinate manifolds would not substantially increase manifold widths. We used these insights to devise a model of hierarchically organized categories, composed of randomly oriented Gaussian manifolds with separable centers (Fig.1E). The model replicated all the features of the IT population responses (Fig.1F-G), providing a viable possibility for the representational geometry of hierarchical categories in the visual cortex.

References

1. E. Rosch, B. B. Lloyd (1978).
2. R. Kiani, H. Esteky, K. Mirpour, K. Tanaka, *Journal of Neurophysiology* **97**, 4296–4309 (2007).
3. N. Kriegeskorte *et al.*, *Neuron* **60**, 1126–1141 (2008).