A Geometric Theory for Complex Cells

Abstract

A fundamental role of the visual cortex is to represent structure in natural scenes, including local features and their smooth transformations. It is widely believed that simple cells learn and recognize these local features, which can be modeled using sparse coding. Previous models of complex cells have been motivated either by the objective to build invariant image representation or to reflect the multivariate higher-order statistics of neural responses. Here we provide a unifying geometric perspective for understanding simple and complex cells based on the recently proposed Sparse Manifold Transform (SMT) model. The SMT is a hierarchical model combining sparse coding (neurons in the first layer) and manifold smoothing (neurons in the second layer) to create temporally smooth representations that reflect transformations in the sensory input. By constraining the second layer in the SMT to have minimum-wiring length, the units in this layer exhibit complex cell properties. Learning in the model leads to cells in layer 1 and 2 which resemble simple and complex cells, respectively. Panel A below shows that orientations and locations of receptive fields (RFs) of layer 1 cells pooled by the same layer 2 cell are similar. Panel B shows this for 21 randomly selected complex cells, visualized as a needle plot representing Gabor fits to each simple cell subunit. During a phase sweep of the optimal grating stimuli for one layer 2 cell, the response of cells in layer 1 is phase sensitive (Panel C, dotted lines) and phase invariant for the layer 2 cell (Panel C, red line). The F1/F0 ratio distribution of all cells in the model is bimodal, separating cells in the two layers (Panel D). Our model provides concise functional explanations for both cell types: Simple cells represent a discretized sampling of the smooth data manifold in the sensor space, while complex cells represent localized smooth functions on the manifold. Thus, the hierarchical network enables complex cells to span the manifold and build an untangled population representation which tends to preserve the identity of the signal while straightening transformations.