

Implicit coding of location, scale and configural information in feedforward hierarchical models of the visual cortex

Cheston Tan ¹,
Thomas Serre²,
Gabriel Kreiman³ and
Tomaso Poggio⁴



 Author Affiliations

Abstract

Feedforward hierarchical models of the visual cortex constitute a popular class of models of object recognition. In these models, position and scale invariant recognition is achieved via selective pooling mechanisms, resulting in units at the top of the hierarchy having large receptive fields that signal the presence of specific image features within their receptive fields, irrespective of scale and location. Hence, it is often assumed that such models are incompatible with data that suggest a representation for configurations between objects or parts. Here, we consider a specific implementation of this class of models (Serre et al, 2005) and show that location, scale and configural information is implicitly encoded by a small population of IT units.

First we show that model IT units agree quantitatively with the coarse location and scale information read out from neurons in macaque IT cortex (Hung et al, 2005). Next, we consider the finding by Biederman et al (VSS 2007) that changes in configuration are reflected both behaviorally and in the BOLD signal measured from adaptation experiments. Model results are qualitatively similar to theirs: for stimuli consisting of two objects, stimuli that differ in location (objects shifted together) evoke similar responses, while stimuli that differ in configuration (object locations swapped) evoke dissimilar responses. Finally, the model replicates psychophysical findings by Hayworth et al. (VSS 2007), further demonstrating sensitivity to configuration. Line drawings of objects were split into complementary pairs A and B by assigning every other vertex to A, and complementary vertices to B. Scrambled versions A' and B' were then generated. Both human subjects and the model rated A as more similar to B than to A'.

Altogether, our results suggest that implicit location, scale and configural information exists in feedforward hierarchical models based on a large dictionary of shape-components with various levels of invariance.

Received April 22, 2008.

© 2008 ARVO