Representational similarity analysis to identify visual action codes for humans, androids and robots

Burcu Aysen Urgen

Abstract:

Visual processing of actions is supported by a network of brain regions in occipito-temporal, parietal, and premotor cortex in the primate brain, known as the Action Observation Network (AON). Although the involvement of these regions in action perception has been well-established, their representational properties remain to be specified. One approach to address this question is to computationally model the visual stimuli that the brain is exposed to, and relate the models of the same stimuli to the brain responses. To this end, we used powerful computational tools from computer vision (CV) as well as attribute-based semantic models to represent the videos consisting of natural actions, and linked those models to the brain responses (fMRI) using representational similarity analysis (RSA), which is ideal for condition-rich experiments (unlike univariate analysis), and inferring representational spaces of the cortex.

RSA technique allows researchers to quantify how similar the neural patterns that correspond to different conditions of an experiment within a certain brain region. Moreover, comparison of the similarity structures across brain regions allows one to study how neural representations change along the cortical hierarchy. In our study, subjects were shown video clips of three agents (a human, an android, and a robot) performing eight different actions during fMRI scanning. We then computed the representational dissimilarity matrices (RDMs) for each brain region, and compared them with that of two sets of model representations that were constructed based on computer vision and semantic attributes.

Our findings reveal that different nodes of the AON have different representational properties. While lower regions of the cortical hierarchy such as early visual cortex represents low level visual features, higher level visual areas such as pSTS represents movement kinematics. Furthermore, as one goes higher in the AON hierarchy, representations become more abstract and semantic as our results revealed that parietal cortex represents several aspects of actions such as action category, intention of the action, and target of the action. These results suggest that during visual processing of actions, pSTS pools information from visual cortex to compute movement kinematics, and passes that information to higher levels of AON coding semantics of actions such as action category, intention of action, and target of action, consistent with computational models of visual action recognition.