in: Proc. Bernstein Conference on Computational Neuroscience (BCCN'09),
Frankfurt, Germany, pp. 36-37, Frontiers in Computational Neuroscience

An Integrated System for Incremental Learning of Multiple Visual Categories

St. Kirstein, H. Wersing, H.-M. Groß, E. Körner

An amazing capability of the human visual system is the ability to learn an enormous repertoire of visual categories. This large amount of categories is acquired incrementally during our life and requires at least partially the direct interaction with a tutor. Inspired by child-like learning we propose an architecture for learning several visual categories in an incremental and interactive fashion based on natural hand-held objects, which typically belong to several different categories. To make the most efficient use of the rare interactively collected training examples a learning method is required which is able to decouple the representation of cooccuring categories. Especially such decoupled representation can not be learned with typical categorization systems so that each category has to be trained independently. This independent training of categories is impractically for interactive learning, because for each category an object belongs to a repetitive presentation to the system is required to train each particular category. We also impose no restrictions to the viewing angle of presented objects, relaxing the common constraint on canonical views. As a consequence this relaxation considerably complicates the category learning task, because in addition to category variations also variations caused by full object rotation has to be handled by the learning method.

The overall categorization system is composed of a figure-ground segregation part and several feature extraction methods providing color and shape features, which for each object view are concatenated into a high-dimensional but sparse feature vector. The major contribution in this paper is an incremental category learning method that combines a learning vector quantization (LVQ) to approach the "stability-plasticity dilemma" with a category-specific forward feature selection to decouple cooccuring categories. Both parts are optimized together to ensure a compact and efficient category representation, which is necessary for fast and interactive learning. Based on this learning method we are able to interactively learn several color (e.g. red, green, blue, yellow and white) and shape categories (e.g. toy car, rubber duck, cell phone, cup, can, bottle, tea box, tools, and four legged animal) with good generalization to previously unseen category members, but also good rejection of unknown categories.

The complete categorization system runs on a single computer, but makes efficient use of currently available multi-core CPUs. Overall the system roughly runs at the frame rate of our current camera system of approximately 6-8 Hz, which is fast enough to show the desired interactive and life-long learning ability. To our knowledge this is the first online learning system which allows category learning based on complex-shaped objects held in hand. Especially the ability

to handle high-dimensional but sparse feature vectors is necessary to allow interactive and incremental learning, where often additional dimension reduction techniques like the principal component analysis (PCA) are required to allow online learning. This high feature dimensionality is also challenging for the used feature selection method, because of the large amount of possible feature candidates. Nevertheless our proposed learning system is able to extract small sets of category-specific features out of many possible feature candidates.