DIRAC

Detection and Identification of Rare Audio-visual Cues

Alon Zweig, Barbara Caputo

Today's computers can do many amazing things but there are still many "trivial" but important tasks they cannot do well. In particular, current information extraction techniques perform well when event types are well represented in the training data but often fail when encountering information-rich unexpected rare events. DIRAC project addresses this crucial machine weakness and aims at contributing theory and algorithms for detection, identification and classification of possibly threatening rare events from the information derived by audio-visual sensors. As a step towards this goal, we present here recent work, developed within the consortium, on object categorization. This is a much investigated and challenging task. In addition to the many parameters influencing the appearance of a single object in the visual domain: lighting, occlusion,...), the task of object class recognition introduces another aspect of complexity, that of intra-class variability. We present two recent approaches for dealing with the categorization task: The first, relying only on visual cues, proposes the use of natural object hierarchy as a means of dealing with object variability and improving object class recognition in a small sample scenario. We identify inherent differences between classifiers trained for different levels of the hierarchy and present a framework for exploiting these differencs by combining classifiers from different levels. Motivated by the observation that biological systems combine several modalities in order to achieve robust categorization, our second approach proposes the use of audio and visual information. We show that combining these cues improves categorization performance and we further investigate this issue by addressing the question of how this combination should be done. Our results suggest that a high-level integration is the most suitable approach, this is consistent with results from neuroscience.