Thesis advisor
Marco Baroni, CIMeC, Università degli Studi di Trento

Author
Adam Liška

Semantic Spaces from Images

Abstract

In vector space models, the meanings of concepts or words are represented as points in high-dimensional vector spaces, also referred to as semantic spaces. These spaces are usually derived automatically from large collections of texts, directly exploiting the sets of contexts in which individual words appear and considering these contexts to be the key constituents of semantic meaning.

The weakness of existing models is that they completely rely on linguistic input, although there is a growing body of evidence that other modalities also contribute to forming semantic representations. This has led to attempts to enrich existing semantic spaces with perceptual information. However, there are two problems that need to be addressed: (1) How to incorporate perceptual information into existing vector spaces? (2) How to acquire this type of data?

In this work, we focus mostly on the latter problem and present a framework for extraction of concept representations from images. The procedure starts with a collection of images that are all tagged with a given concept. Each image is represented by a set of features called visual words. These features constitute the visual context of the given concept similarly to the linguistic context used in textual models. In this way, we arrive at having two sets of contexts for each concept: a set of linguitic contexts derived from text corpora and a set of visual contexts extracted from images. These two sources of information are then used to create two distinct conceptual representations, textual and visual, which can be combined to create a final representation of the given concept.

Key words: distributional semantics, lexical semantics, semantic memory, concept representation