Chapter 29
A Generic Framework for Feature Representations in Image Categorization Tasks

Adam Csapo
Budapest University of Technology and Economics, Hungary

Barna Resko
Hungarian Academy of Sciences, Hungary

Morten Lind
NTNU, Dept. of Production and Quality Engineering, Norway

Peter Baranyi
Budapest University of Technology and Economics, Hungary, & Hungarian Academy of Sciences, Hungary

Domonkos Tikk
Budapest University of Technology and Economics, Hungary

ABSTRACT

The computerized modeling of cognitive visual information has been a research field of great interest in the past several decades. The research field is interesting not only from a biological perspective, but also from an engineering point of view when systems are developed that aim to achieve similar goals as biological cognitive systems. This paper introduces a general framework for the extraction and systematic storage of low-level visual features. The applicability of the framework is investigated in both unstructured and highly structured environments. In a first experiment, a linear categorization algorithm originally developed for the classification of text documents is used to classify natural images taken from the Caltech 101 database. In a second experiment, the framework is used to provide an automatically guided vehicle with obstacle detection and auto-positioning functionalities in highly structured environments. Results demonstrate that the model is highly applicable in structured environments, and also shows promising results in certain cases when used in unstructured environments.

DOI: 10.4018/978-1-4666-0261-8.ch029

The research areas inspired by cognitive science have significantly broadened and have received more attention in the last few years. Fields in engineering dealing with high-complexity intelligent systems tend to build more often on the results of cognitive science. One such example is the growing popularity of cognitive informatics based tools in modern image processing, object recognition and machine vision, designed with the motivation of cognitive science.

The algorithmic difficulties arising in computerized object recognition have been evident since the 1960s. Due to the realization that information sciences alone cannot successfully tackle the problem of object recognition, research in cognitive psychology, brain physiology and artificial intelligence have advanced hand in hand to better understand human visual processing. I. Biederman's results in the late 1980's (Biederman, 1987) - proving the indispensability of low-level visual features such as vertices and corners in the successful recognition of objects - have sparked increased interest in developing algorithms that make use of such geometrical primitives (some examples are D. Lowe's SIFT method (Lowe, 1999), A. Berg's geometric blur operator (Berg & Malik, 2001), and various works of M. Riesenhuber and T. Poggio (Riesenhuber & Poggio, 1999, 1997; Serre, Wolf, Bileschi & Riesenhuber, 2007).

Due to the increasing relative importance of cognitive inspired methods, it is important to model the processes described by cognitive science and to integrate such functions in a toolbox of cognitive informatics. This paper addresses this need by providing a general computational model of low-level feature extraction for visual information processing. The computational model - besides providing a generalized concept of visual features - allows for a highly differentiated but at the same time uniform representation of visual features. The operations defined on the visual features are all linear operations and have the advantage of requiring extremely low computational power.

The applicability of the model is demonstrated through two examples. In the first application, the proposed framework is used to generate input data structures to a linear classifier that has been shown to perform powerfully in the categorization of text documents. The method is evaluated on natural images taken from the Caltech 101 database, and is shown to yield promising results. In the second application, the representational framework is used to help automatically guided vehicles to autonomously self-locate and detect obstacles within highly structured environments.

Although the proposed framework treats all parts of images in a uniform way and does not rely on explicitly pre-determined keywords, it can be used in combination with other widespread methods within cognitive informatics, such as keyblock-based image retrieval methods (Zhu, Rao & Zhang, 2002), and various bag-of-keypoints techniques (Csurka, Dance, Fan, Willamowski & Bray, 2004; Grauman & Darrell, 2005; Lazebnik, Schmid & Ponce, 2006). In the examples given in this paper, however, we prefer to use the most naive approaches possible to demonstrate the viability of our model in the most general cases.

THE VISUAL FEATURE ARRAY

The Visual Feature Array Concept

At the heart of the proposed concept is a Visual Feature Array (VFA) model, which is a cognitive information processing model that uses the information processing structures defined in the VFA concept. The VFA model obtains information from the environment, performs various operations on it, and then supplies higher-order models of cognitive informatics with its output. The proposed VFA concept allows for the implementation of cognitive functions analogous to those performed in the primary visual cortex (Figure 1).