

Object categorisations using templates constructed from multi-scale line and edge representations

S. Nunes, D. Almeida, J. Rodrigues, J.M.H. du Buf
University of Algarve, Faculty of Sciences and Technology,
8000-810 Faro, Portugal
{srnunes,dalmeida,jrodrig,dubuf}@ualg.pt
<http://w3.ualg.pt/~dubuf/vision.html>

Object categorisation is linked to detection, segregation and recognition. In the visual system, these processes are achieved in the ventral “what” and dorsal “where” pathways [3], with bottom-up feature extractions in areas V1, V2, V4 and IT (what) in parallel with top-down attention from PP via MT to V2 and V1 (where). The latter is steered by object templates in memory, i.e. in prefrontal cortex with a what component in PF46v and a where component in PF46d.

We are studying three related problems: when, where and how does categorisation take place. The “when” problem allows for two hypotheses. The easy one is to assume that categorisation occurs after recognition [4]: if specific neurons respond in the case of recognising dog-1, dog-2 and dog-3, a grouping cell can combine responses: a dog. This view is too simplistic, because the system must collect evidence for a specific object or object group in order to select possible templates in memory. For example, when we glance a painting of Arcimbaldo, the 16th-century Italian painter, our first reaction is “a face!” but then follows “fruits?” and finally “the cheek is an apple!”

When categorisation occurs before recognition, the “where” problem is, at least partially, solved: it must take place at a very high level, with access to object templates, and just before recognition. In fact, recognition can be seen as a last categorisation step. Therefore, the “how” problem can be solved by taking into account feature extractions in V1 and beyond and the propagation of features to higher cortical areas. During the last years, we concentrated on the extraction of low-level primitives: lines, edges and keypoints, all multi-scale [5]. Keypoint scale space provides ideal information for constructing saliency maps for focus-of-attention, and the grouping of keypoints at different scales is robust for e.g. face detection [6]. However, detection in the where pathway must be linked with categorisation and recognition in the what pathway (a face, whose face?). The balance between the use of lines/edges and keypoints in the two pathways remains an open question.

A less open question concerns the balance between features detected at different scales: information at coarse scales propagates first to higher layers, after which information at progressively finer scales arrives there [2]. This probably implies that coarse-scale information is used for a first but rapid categorisation, after which categorisation is refined until the object is recognised. It has been proposed that a first categorisation is based on a lowpass filtered image of the object [1], but a smeared blob lacks structure. In our own experiments we therefore applied a different approach: after segregation, the coarse-scale line/edge representation of the outline is used for pre-categorisation, after which all information is used for final categorisation. We studied four groups [7]: cars, mugs, fruits and animals. The latter groups consisted of apples, pears and tomatoes; and of horses, cows and dogs. Pre-categorisation of the four groups was perfect. Problems in the final categorisation were due to not using colour (apples vs tomatoes) and minute differences between dogs and horses.

- [1] Bar, M. (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neuroscience* 15, pp. 600-609.
- [2] Bar, M. (2004) Visual objects in context. *Nature Reviews: Neuroscience* 5, pp. 619-629.
- [3] Deco, G., Rolls, E. (2004) A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Res.* 44, pp. 621-642.
- [4] Riesenhuber, M., Poggio, T. (2000) Cbf: A new framework for object categorization in cortex. 1st IEEE Int. Worksh. Biologically Motivated Computer Vision, Seoul (Korea), pp. 1-9.
- [5] Rodrigues, J., Almeida, D., Nunes, S., Lam R. and du Buf, J.M.H. (2005) Building the what and where systems: multi-scale lines, edges and keypoints. Worksh. Active Vision (7WAV2005), Univ. of Reading, UK, Sept. 12, p. 8.
- [6] Rodrigues, J. and du Buf, J.M.H. (2006) Multi-scale keypoints in V1 and beyond: object segregation, scale selection, saliency maps and face detection. Accepted for: *BioSystems* (Elsevier).
- [7] Rodrigues, J. and du Buf, J.M.H. (2006) Cortical object segregation and categorization by multi-scale line and edge coding. *Int. Conf. Computer Vision, Theory and Applications (VISAPP2006)*, Setubal, Portugal, Feb. 25-28, Vol. 2, pp. 5-12.

[Supported by FCT program POSI, framework QCA III, and by PRODEP III Medida 5-Acção 5.3.]