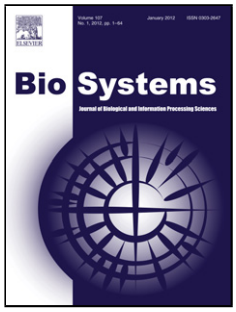

Title: Proto-Object Categorisation and Local Gist Vision using Low-Level Spatial Features

Author: Jaime A. Martins J.M.F. Rodrigues J.M.H. du Buf

PII: S0303-2647(15)00094-5
DOI: http://dx.doi.org/doi:10.1016/j.biosystems.2015.07.001
Reference: BIO 3583

To appear in: *BioSystems*

Received date: 1-5-2014
Revised date: 2-7-2015
Accepted date: 2-7-2015

Please cite this article as: Jaime A. Martins, J.M.F. Rodrigues, J.M.H. du Buf, Proto-Object Categorisation and Local Gist Vision using Low-Level Spatial Features, *BioSystems* (2015), http://dx.doi.org/10.1016/j.biosystems.2015.07.001

Proto-Object Categorisation and Local Gist Vision using Low-Level Spatial Features

Jaime A. Martins[a,c], J. M. F. Rodrigues[b,c], J. M. H. du Buf[a,c]

[a] *Vision Laboratory (FCT), LARSyS*
[b] *Vision Laboratory (ISE), LARSyS*
[c] *University of the Algarve, Portugal*

Abstract

Object categorisation is a research area with significant challenges, especially in conditions with bad lighting, occlusions, different poses and similar objects. This makes systems that rely on precise information unable to perform efficiently, like a robotic arm that needs to know which objects it can reach. We propose a biologically-inspired object detection and categorisation framework that relies on robust low-level object shape. Using only edge conspicuity and disparity features for scene figure-ground segregation and object categorisation, a trained neural network classifier can quickly categorise broad object families and consequently bootstrap a low-level scene gist system. We argue that similar processing is possibly located in the parietal pathway leading to the LIP cortex and, via areas V5/MT and MST, providing useful information to the Superior Colliculus for eye and head control.

Keywords: Disparity, 3D, stereo vision, colour, population coding, learning, biological model, figure-ground, segregation, object, categorisation, verification, neural network, visual cortex.

1. Introduction

There are many visual pathways related to object categorisation and recognition, especially focusing on quick shape categorisation, which is essential for scene gist. Obvious information sources for object shape and segregation are colour, texture, motion and depth from stereo. However, much richer information is available. Apart from depth from stereo, velocity gradients of optical flow can be used to locally encode ordinal depth at surface borders and also, but globally, ego-motion [1]. In addition, in case of occlusions, figure-ground segregation by local (and often intricate) border-ownership relations of e. g. vertex (keypoint) structures is also possible, as hypothesised in the discussion paper by Kogo and Wagemans [2], which attracted many comments. How our visual system extracts and integrates all information is still rather speculative.

In this paper we focus on the transition between low-level syntax and low-level semantics, using elementary information such as surface lighting, colour and stereo disparity. The goal is to develop an integrated system for fast local gist vision: *which types* of objects are *about where* in a scene. This is necessary to bootstrap and guide, even alleviate, the processing in the ventral and dorsal data streams. These streams are known to serve two goals: the dorsal stream, also called the *where* or *vision-for-action* stream, is mostly devoted to optical flow and stereo disparity, whereas the ventral stream, also called the *what* or *vision-for-perception* stream, is devoted to object categorisation and recognition [3, 4]. However, the dorsal stream can also play a very important role in fast object categorisation [5, 6, 3], which is the main focus of this paper.

An integrated system must first solve two hard problems: (a) The first one is of paradoxical nature, as precise object categorisation and recognition in the ventral stream requires object segregation, but object segregation has usually been regarded only possible if the system already knows what the object is (assuming of course that objects are complex and that they are seen against equally complex backgrounds). Consequently, we explore the possibility

Email addresses: jamartins@ualg.pt (Jaime A. Martins), jrodrig@ualg.pt (J. M. F. Rodrigues), dubuf@ualg.pt (J. M. H. du Buf)