J Sign Process Syst
DOI 10.1007/s11265-009-0349-0

Markov Random Fields with Asymmetric Interactions for Modelling Spatial Context in Structured Scene Labelling

Daniel Heesch · Maria Petrou

Received: 6 February 2008 / Revised: 29 January 2009 / Accepted: 2 February 2009

Abstract In this paper we propose a Markov random field with asymmetric Markov parameters to model the spatial and topological relationships between objects in structured scenes. The field is formulated in terms of conditional probabilities learnt from a set of training images. A locally consistent labelling of new scenes is achieved by relaxing the Markov random field directly using these conditional probabilities. We evaluate our model on a varied collection of several hundred hand-segmented images of buildings. The incorporation of spatial information is shown to improve greatly the performance of some trivial classifiers.

Keywords MRF · Scene labelling · Random field relaxation · Machine learning · Contextual vision · Spatial context

1 Introduction

Recent years have seen notable improvements in the performance of object classifiers. Greater robustness against occlusion and intraclass variability has been achieved by describing objects by a large number of local and largely view-invariant features (e.g. [6, 19, 20, 23]). For a specific class (e.g. face detection), efficient classification methods, such as boosting, allow recognition to be in real-time (e.g. [22]). Some of these models have the additional benefit of biological plausibility. The hierarchical feed-forward architecture of [18], for example, aims to mimic the ventral stream of visual information processing and is able to predict with great accuracy whether or not an object is present in a scene.

It seems, however, that, in order to be able to scale to the several thousands of categories humans discriminate without effort, appearance based object classification needs to be complemented by techniques that utilise contextual information. Context may be described as any dependency between the object to be recognised and everything else in the scene, be this other objects or the scene as a whole. Experimental evidence suggests that humans do exploit both types of dependency during object recognition. It is well established, for example, that the nature of a scene can be recognised based on low spatial frequency information [13]. Recent neuro-imaging studies support the view that low spatial frequencies are processed in the cortex at a very early stage during visual recognition [2]. Spatial relationships between objects are one example of contextual information that is captured by the low frequency content of an image. In fact, Bar and Aminoff in [1] establish "early activation of cortical context networks" that appear to store spatial relationships. Spatial context may thus play a key role as an early facilitator during object recognition.

The purpose of this work is to demonstrate the importance of contextual vision over vision based on the use of individual characteristics, using a scaled down version of a visual recognition problem. Our goal is to learn these spatial and topological relationships from a set of hand-labelled data and to utilise this information in a Markov random field (MRF) model to achieve a consistent labelling of new scenes. The MRF is defined

D. Heesch · M. Petrou (✉)
Communications and Signal Processing Group,
Electrical and Electronic Engineering, Imperial College
London, Exhibition Road, London SW7 2AZ, UK
e-mail: maria.petrou@imperial.ac.uk

Published online: 24 February 2009