Computer Vision and Image Understanding 122 (2014) 47–64

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Biologically-inspired robust motion segmentation using mutual information

Anna-Louise Ellis, James Ferryman *

Computational Vision Group, School of Systems Engineering, University of Reading, Whiteknights, Reading, Berkshire RG6 6AY, UK

ARTICLE INFO

Article history:
Received 15 April 2013
Accepted 27 January 2014

Keywords:
Biologically-inspired vision
Background modelling
Segmentation
Surveillance
Performance evaluation

ABSTRACT

This paper presents a neuroscience inspired information theoretic approach to motion segmentation. Robust motion segmentation represents a fundamental first stage in many surveillance tasks. As an alternative to widely adopted individual segmentation approaches, which are challenged in different ways by imagery exhibiting a wide range of environmental variation and irrelevant motion, this paper presents a new biologically-inspired approach which computes the multivariate mutual information between multiple complementary motion segmentation outputs. Performance evaluation across a range of datasets and against competing segmentation methods demonstrates robust performance.

1. Introduction

The ability to extract objects of interest from video sequences, using detected motion, remains an active area of research within the computer vision community. The capacity to provide real-time segmentations – silhouettes and bounding boxes – of objects (especially pedestrian) assists in the tracking and reasoning of the behaviour. Surveillance scenes often contain change that may be inaccurately detected as object motion such as changes in lighting, periodic motion, moving shadows and reflections. In addition the quality of surveillance footage is often poor, and at a low resolution resulting in noisy motion and ghosts. An example of these challenges is shown in Fig. 1. The extraction of objects of interest is frequently tackled by removing all irrelevant pixels in each frame. This is referred to as motion segmentation. To date no segmentation algorithm is robust under all these conditions.

In this paper, we propose a new formulation of pixel-based foreground segmentation which is motivated by recent results in biological vision which exploit the mutual information between multiple segmentation channels. The paper is divided as follows. Firstly, Section 2 details the biological motivation and mapping to a combination of parametric background modelling approaches. This is followed in Section 3 by approaches to fusing the outputs of multiple segmentation algorithms and introduces the multivariate mutual information forumulation adopted in this work. In Section 4 the datasets, evaluation methodology and the results of experiments are presented before concluding in Section 5 with conclusions and recommendations for future research.

2. Biologically-inspired segmentation

The ability of primates to recognise objects of interest, regardless of illumination and background, drives much of the biologically inspired computational vision systems. A new biologically inspired vision system is introduced in this section that models current vision research which has not previously been examined by the computational vision community.

In Section 2.1 the model of primate vision conventionally accepted by the computer vision community is presented. Section 2.2 provides descriptions of state of the art biologically inspired computational vision systems that refer to this model. Section 2.3 progresses onto accounts of current published neuro-biological, physiological and psychological vision research and highlights descriptions of retinal functions, inputs to the ventral and dorsal streams, and ventral and dorsal stream behaviour that have not been considered in modelling primate visual systems in the computer vision community. Based on this, a new model of understanding is presented and the behaviours of these retinal functions are summarised.

2.1. Conventional model of primate vision

It is widely acknowledged that the rods and cones (photoreceptors) of the primate retina detect light and cells of the inner retina providing the initial stages of the visual processing. The retinal

* Corresponding author. Fax: +44 118 9751994.
E-mail address: james@computer.org (J. Ferryman).

http://dx.doi.org/10.1016/j.cviu.2014.01.009
1077-3142/

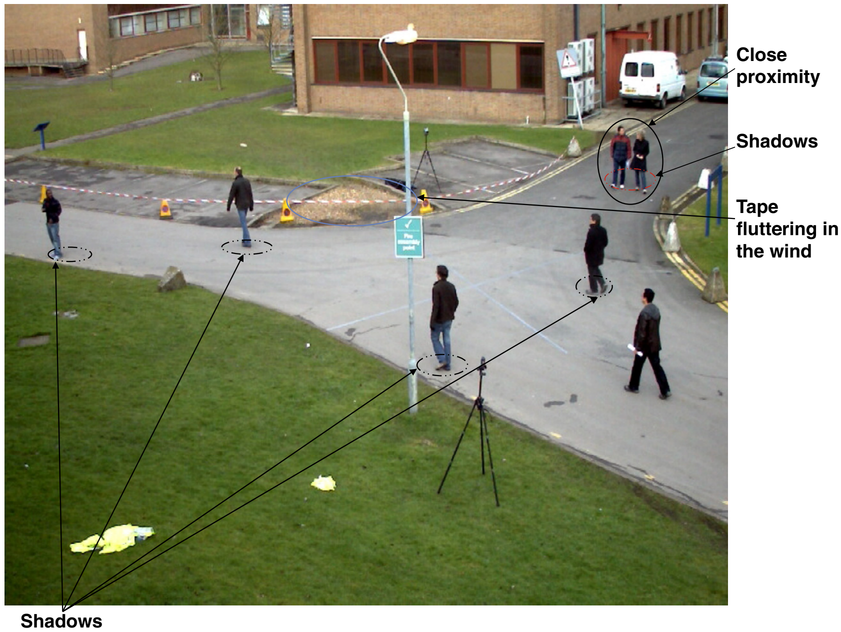

Fig. 1. PETS 2009 dataset original frame annotated with automated visual surveillance challenges.

ganglion cells convey this information, via pathways in the lateral geniculate nucleus, to the ventral and dorsal streams in visual cortex. Fig. 2 represents a model of these traditionally accepted components, frequently referred to in biologically inspired computational vision systems.

Within the retina, shown in Fig. 2 as the blue area, the photoreceptor rod cells respond to achromatic brightness and the photoreceptor cone cells respond to short (blue), medium (green) and long (red) chromatic wavelengths. These nerve impulses are passed onto the network of horizontal, amacrine and bipolar cells, which provide cumulative information to retinal ganglion cells, shown in Fig. 2 as the midget and parasol ganglion cells. The midget ganglion cells have been associated with providing chromatic information and parasol ganglion cells with luminance and contrast.

The lateral geniculate nucleus (LGN), illustrated as the green area in Fig. 2, receives the assembled information from the ganglion cells, in the form of pathways. The parvocellular pathway is conventionally understood to receive information from the midget ganglion cells, and as such provides a means to direct colour information to the visual cortex. It is customary to describe the magnocellular pathway as a swiftly responsive structure, presenting the visual cortex with luminance and contrast information.

Finally, the visual cortex (VC), emphasised as the purple area in Fig. 2, includes two different streams: the ventral stream, associated with form, and the dorsal stream associated with motion.

2.2. Existing bio-inspired computational models

Ref. [43] state that because bio-inspired vision models based on a vertebrates visual system are limited and require high computational cost, real-time applications are seldom addressed. As flies are capable of exploiting optical flow, which modelled by calculating the local

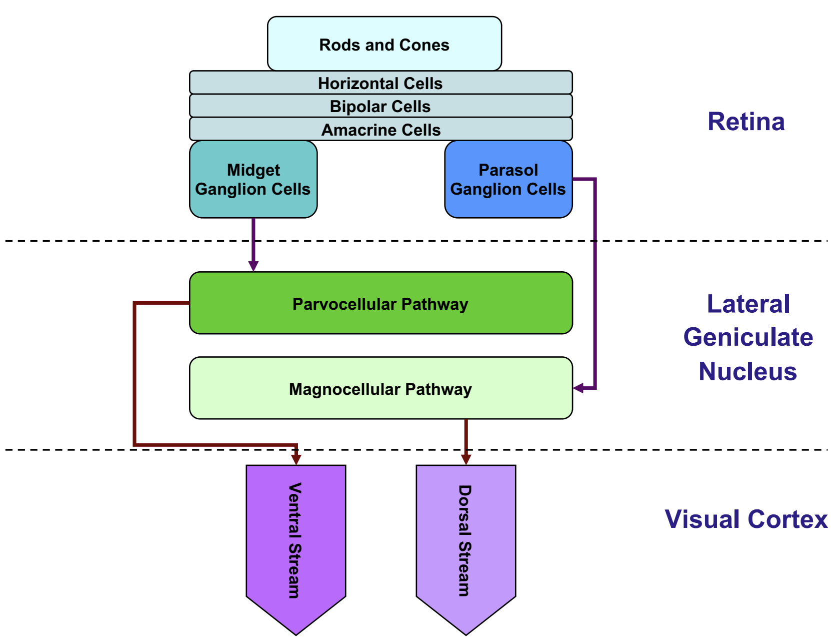

Fig. 2. Model of traditional computational vision process.