Computer Vision and Image Understanding 114 (2010) 758–773

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Using Human Visual System modeling for bio-inspired low level image processing

A. Benoit [a,*], A. Caplier [b], B. Durette [b], J. Herault [b]

[a] LISTIC – Polytech'Savoie, B.P. 80439, 74944 Annecy le Vieux Cedex, France
[b] Gipsa-lab, 961 rue de la Houille Blanche, Domaine Universitaire – B.P. 46, F – 38402 Saint Martin d'Hères Cedex, France

ARTICLE INFO

Article history:
Received 4 December 2008
Accepted 31 January 2010
Available online 4 March 2010

Keywords:
Retina model
V1 cortex model
Low level image processing
Real-time processing
Contour analysis
Motion analysis
Preprocessing
Texture and structure enhancement
Local adaptation

ABSTRACT

An efficient modeling of the processing occurring at retina level and in the V1 visual cortex has been proposed in [1,2]. The aim of the paper is to show the advantages of using such a modeling in order to develop efficient and fast bio-inspired modules for low level image processing.

At the retina level, a spatio-temporal filtering ensures accurate structuring of video data (noise and illumination variation removal, static and dynamic contour enhancement). In the V1 cortex, a frequency and orientation based analysis is performed.

The combined use of retina and V1 cortex modeling allows the development of low level image processing modules for contour enhancement, for moving contour extraction, for motion analysis and for motion event detection. Each module is described and its performances are evaluated.

The retina model has been integrated into a real-time C/C++ optimized program which is also presented in this paper with the derived computer vision tools.

1. Introduction

In this paper, we propose an image processing approach belonging to what we call "biological vision based approach". The basic idea is to copy the Human Visual System (HVS) by modeling some of its parts in order to develop low level image processing modules. Up to now, the most well-known parts of our visual system are the retina and the V1 cortex area which are the two parts on which we focus our work. The retina can be considered as a preprocessing step which conditions the visual data for facilitated high level analysis. The V1 cortex can be considered as a low level visual information describer. From these two "tools", we want to show how to design efficient low level image processing tools.

Biologically inspired methods for image processing are numerous and we choose to focus only on bio-inspired models dedicated to image processing in order to make the paper easier to read. For example, the Retinex filter proposed in [3,4] is a method that enhances a digital image in terms of dynamic range compression, color independence from the spectral distribution of the scene illumination, and color/lightness rendering as it is done in the retina and in the cortex. This algorithm is based on luminance analysis and its enhancement. It assumes that color perception is related to ratios of reflected light intensity in specific wavelength bands computed between adjacent areas. As a consequence, this algorithm is dedicated to color applications. Other models of the HVS are used, for example, for information coding [5]. These methods generally use high level information processing such as visual cortex modeling but do not take into account the low level processing that occurs at retina level.

Since our goal is to demonstrate the interest of using retina and V1 cortex modeling in order to proceed to low level image processing, the preliminary step of our work, which is to choose the most appropriate retina and cortex models, is described in the following. As discussed in [51], the definition of standard models is a rich research field and some approaches allow dedicated image processing implementations to be expected. As far as retina models are concerned, some have already been proposed with different degrees of precision. Mead and Mahowold [6] was a precursor for the modeling of the neurophysiological properties of vertebrate's retinas by considering analogies with electronic circuits. His model focuses on the link between the retinal architecture and its functionalities. Nevertheless, his work insists more on the spatial filtering properties of the retina than on temporal effects related to motion analysis. The modeling of the biological retina was also studied by Franceschini et al. [7] who worked on the retina architecture of the fly. He built robots working on the same model and showed their properties for target tracking or for flying in unstable wind conditions and for collision prevention. Spike based models

* Corresponding author at: Fax: +33 450 09 65 59.
E-mail addresses: alexandre.benoit@univ-savoie.fr (A. Benoit), alice.caplier@gipsa-lab.grenoble-inp.fr (A. Caplier), barthelemy.durette@gipsa-lab.grenoble-inp.fr (B. Durette), jeanny.herault@gipsa-lab.grenoble-inp.fr (J. Herault).
URLs: http://www.listic.univ-savoie.fr (A. Benoit), http://www.gipsa-lab.inpg.fr (B. Durette).

1077-3142/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2010.01.011

were also studied; an advanced model was presented with the SpikeNet toolbox [8]. It models the electrical impulse spikes exchanged by the neural cells at the retina ganglion cells and V1 cortex levels. It already demonstrates high speed computing properties for high level image analysis, but low level retina processing are not completely described. Other approaches are developed such as digital retinas. Some of them are methods dedicated to VLSI (very-large-scale integration) implementations [9,10]. These algorithms are efficient parallel methods generating binary or floating point output pictures but the models contain only parts of all the processes carried out in the retina.

The starting point of our work is an accurate model of the human retina. This model presents a global approach of the retina processing inspired from an analogy between electronic circuits and signal processing strategies of the biologic retina. It is based on Mead's work and has been improved in terms of spatial and temporal properties by Herault and Beaudot [1,2,11]. It describes the different computing carried out in the first cell layers of the retina (Outer and Inner Plexiform Layers). This model allows fine perception modeling. This emphasizes the different cell network properties of the retina and its implementation enables fast computing thanks to natural parallel processing properties.

Considering V1 cortex, several studies led to the creation of various models. Marcelja [12] showed that the cortical cells in the V1 cortex are sensitive to orientations and can be modeled with 1D Gabor filters. This work was extended to 2D by Daugman [13]. This modeling leads to a simple representation of scene information in the spectral domain. In this way, 2D Gabor filters are generally used in literature for texture classification [14], or saliency area research to extract relevant features in a scene [15]. Because of their properties in log scale (reliable zoom effects handling), we propose to use the modeling of the V1 cortex area described in [16] which uses log polar Gabor filters (GloP) instead of Gabor filters.

With the choice of Herault's model for retina modeling and the choice of Guyader's model for the V1 cortex modeling, we obtain a model for the parts of the visual system we are interested in. In order to situate the chosen global HVS model with regard to well known visual system models, we present the main orientations of these works and ours. The Itti and Koch model [17] focuses on the analysis of the visual scene in terms of scale and orientation description. This model exhibits the high level analysis achieved at the visual cortex level in order to compute saliency maps for visual attention modeling. These bottom-up orientation and scale description are indeed specific features of the V1 cortex area which we also propose to perform with the help of Guyader's model [16]. The work of Walter [18] also insists on the processing carried out at the cortex level and adds top-down interactions for visual attention simulation. Nevertheless, at the retina level, low level processing is not fully considered. Similar approaches have been proposed, for example by Daly's [19], the Irccyn Lab's model [20] and Gipsa Lab's model [63]. These models are suited for image and video quality evaluation and saliency area extractions. These accurate models insist more on high level cortex processing (even above V1 area) dedicated to image description than on the properties of low level processing done at the retina level. The Contrast Sensitivity Function (CSF) they use does not include some specific features of the retina such as local adaptation and temporal filtering. In comparison, our approach focuses more on the first low level retina processing and the V1 cortex in the aim of demonstrating the interest of the low level retina filtering properties. Future work will consist in fusing our model with aforementioned approaches in order to reach a higher step of complexity with a more accurate low level processing precision and to describe a wider area of image processing applications.

In order to show the potential of such human visual models for efficient low level image computing, we present, in this paper, a set of real-time image processing modules. A first set, based on a retina model, allows detail and motion information extraction. A second set based on the V1 cortex area and a motion event detector enable to describe the visual scene at a higher semantic level. Keep in mind that the motion analysis which will be exposed can be considered as being close to optical flow computation. Nevertheless, we insist more on the preprocessing aspect of the retina for motion energy extraction, its noise reduction and local motion information enhancement. Even if information about motion energy is offered, the extraction of the optical flow is the next step, as proposed in [21].

The paper is presented as follows: Section 2 gives a short description of the model proposed in [1,2] for retina and V1 cortex processes. Section 3 describes the four low level processing modules which we developed for contour enhancements, moving contour extraction, image orientation analysis and context aware motion event detection. Section 4 describes the developed real-time retina program which has been made available publicly and which justifies our global approach, consisting in using HVS modeling in order to build up efficient image processing algorithms. This section also summarizes the previous image processing algorithm which has already been achieved with the help of the presented models.

2. Human Visual System modeling

Fig. 1 gives a general overview of the parts of the HVS which are considered here and which have been modeled in [1,2,16]. In the retina, the spatial and temporal properties of the different cells layers are considered, from photoreceptors and the connected cell layers of the so called Outer Plexiform Layer (OPL) followed by the Inner Plexiform Layer (IPL). These processing steps are described in Section 2.1. The proposed model allows two information channels to be modeled. The former, Parvo, being related to details extraction while the latter, Magno, is dedicated to motion analysis. In the V1 cortex area (cf. Section 2.2), a frequency and orientation analysis in the log polar domain is carried out [16]. It is intended to process the information given by the retina model in order to perform low level visual scene properties description.

2.1. Retina modeling

Fig. 2 describes the different retina cells: photoreceptors, horizontal cells, bipolar cells, ganglion cells and amacrine cells. Photoreceptors are responsible for visual data acquisition and are also associated with a local logarithmic compression of the image lumi-

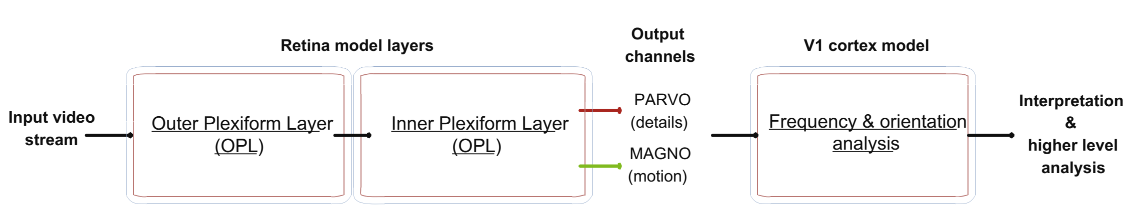

Fig. 1. General algorithm.