Biologically Inspired Cognitive Architectures (2014) 7, 26–38

Available at www.sciencedirect.com

ScienceDirect

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

RESEARCH ARTICLE

Visual attention model for computer vision

F. Robert-Inacio [a,b,*], L. Yushchenko [b]

[a] CNRS Institut Matériaux Microélectronique et Nanosciences de Provence, IM2NP UMR 7334, Place Pompidou, 83000 Toulon, France
[b] Institut Supérieur de l'Électronique et du Numérique de Toulon, IM2NP, 83000 Toulon, France

Received 9 August 2013; received in revised form 4 November 2013; accepted 5 November 2013

KEYWORDS
Radial sampling;
Focus of attention;
Biologically inspired pavement;
Data compression;
Foveated image;
Eye saccades

Abstract
In order to model some abilities of the human eye, it is very useful to pay more attention to the visual system organization. In this way the focus of attention principle can be directly described from cone processing when acquiring color images. Cones are photoreceptor cells located in the retina and with the highest density in a small area called fovea. These cells are sensitive to well-defined light wavelengths and are at the basis of color perception. But cones are non-uniformly located over the retina. They are distributed in a radial way from the fovea center. That is why it is necessary to define radial sampling according to a focusing point in order to reconstruct images as they are supposed to be captured by the retina. Such a radial sampling enables as well to obtain compressed data with a large ratio. Although compression is achieved with loss, global information is preserved on the whole image (foveated image). In this paper, the hexagonal cell model is selected to achieve eye saccade mimicking because hexagonal cells are the closest in shape to retinal cells. Furthermore, this cell model gives the best results in terms of data preservation. Such a sampling leads to setting up foveated image processing. In this way, image processing is achieved on less data (15 times less) and so performed in a really faster way. Furthermore foveated images are also used to compute a sequence of points of interest. By following this sequence, a system of vision can mimick eye saccades when focusing successively at each point of interest. This process is a part of the whole process modeling visual attention, as it takes into account detection of points of interest.

1. Introduction

When exploring a new scene, the human visual system computes several points of interest in order to focus at them

* Corresponding author at: CNRS Institut Matériaux Microélectronique et Nanosciences de Provence, (IM2NP UMR 7334), Place Pompidou, 83000 Toulon, France. Tel.: +33 494038997.
E-mail address: frederique.robert@im2np.fr (F. Robert-Inacio).

2212-683X/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bica.2013.11.001

(eye saccades). Simulating visual attention and mimicking eye saccades are very useful processes in computer vision. Visual attention can allow object recognition in a scene (Posner & Fan, 2007; Walther, Rutishauser, Koch, & Perona, 2005). Visual attention can also be studied in the detection of interesting points on videos (Zhai & Shah, 2006). Applications of such a technique are numerous, for example in the medical field, to provide prosthesis to blind or quasi-blind people (Sawan, Gosselin, & Coulombe, 2008), in cognitive studies, to determine the human behavior when facing a new scene (Cornelissen, Bruin, & Kooijman, 2005), in computer vision to validate methods by comparing them to a human observer abilities (Chen et al., 2001), and so on.

A set of focused images is the database for the visual cortex to reconstruct the whole scene (Liversedge, Gilchrist, & Everling, 2011). The focus of attention principle can be modelized by considering a particular point called the focusing point, and by computing a re-sampled image defining the scene as it might have been captured by the retina. There exist two different kinds of cells in the retina: cones and rods. The cone distribution is the model basis because they are involved in photopic (diurnal) vision and in color perception Gibson (1950) and Marr (1982). Cones are photoreceptors acquiring light data and they are located in a very large number in a small retina area called the fovea (Fig. 1) Wandell (1995). Outside the fovea their number decreases rapidly according to the distance to the fovea center. This is linked to visual acuity decreasing from the center of gaze to the temporal or nasal fields (Fig. 2) Levine (1985). Furthermore the fovea center corresponds to the focusing point.

Focus of attention enables then to re-sample images in a radial way (polar space). This can be useful firstly because it allows to decrease considerably the amount of data to be processed and secondly it reproduces the blurring phenomenon associated with visual acuity (Khan, Dinet, & Konik, 2011). These images are encoded in a specific way that

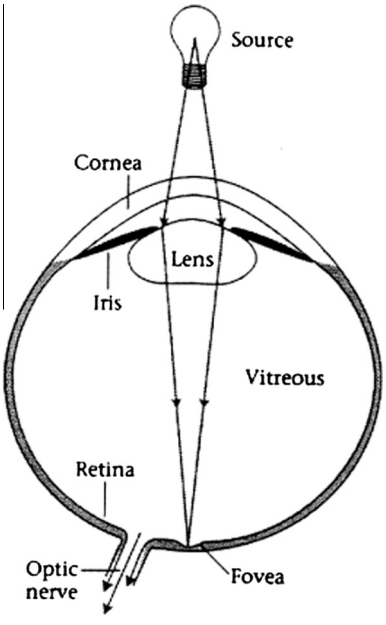

Fig. 1 Eye scheme.

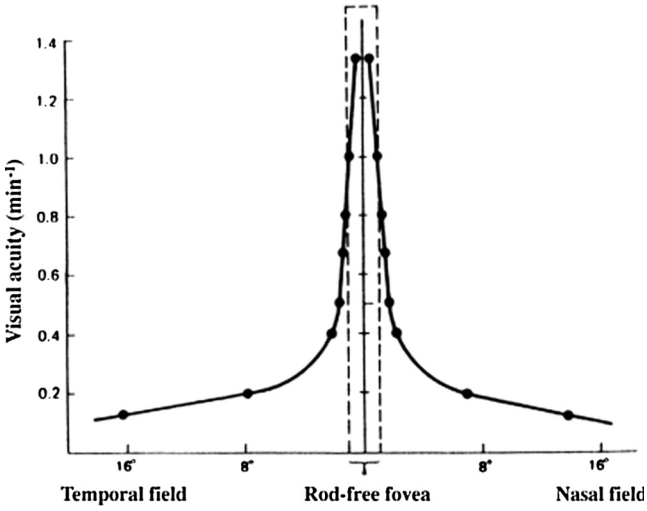

Fig. 2 Visual acuity according to the visual angle.

requires that image processing algorithms must be designed by taking into account the image features.

Then a sequence of focusing points is computed in order to mimick eye saccades (Hernandez, Levitan, Banks, & Schor, 2008). This sequence of focusing points is composed of points of interest directly determined on resampled images.

1.1. Related works

In recent years sensors as well as computers have reached very high performance and capacity. That is why it is now possible to set up active vision systems inspired by the biological scheme corresponding to the human eye and visual cortex (Bailey & Bouganis, 2008). Systems using visual attention are based on schemes from cognitive sciences (Snowden, Thompson, & Troscianko, 2006). Neurobiologists and psychologists have widely studied this topic in collaboration with computer scientists (Frintrop, Rome, & Christensen, 2010). In robotics, modeling visual attention is used to solve real-life problems (Moeslung & Granum, 2001; Vikram, Tscherepanow, & Wrede, 2011). As well applications for smartphones can be designed for automatic resizing of images. Furthermore, the foveation principle which is based on visual attention is also used for video compression (Itti, 2004).

Several studies about eye saccades are currently under consideration. For example, forces required to move the eye can be taken into account in order to simulate eye scanpaths, as well as saliency (Wang et al., 2011). And eye-tracking is a very useful tool to determine eye behavior (Grossberg, Srihasam, & Bullock, 2012; Ntouskos, Pirri, Pizzoli, Sinha, & Cafaro, 2013) and compare it to simulations.

1.2. Contents

In this paper, we propose an original way of sampling images in order to simulate image acquisition by retina. This sampling enables to considerably decrease the amount of data to process. Sampled images can then be used to mimick