Chapter 22

Implementation and Evaluation of a Computational Model of Attention for Computer Vision

Matthieu Perreira Da Silva
IRCCyN – University of Nantes, France

Vincent Courboulay
L3i – University of La Rochelle, France

ABSTRACT

In the field of scene analysis for computer vision, a trade-off must be found between the quality of the results expected and the amount of computer resources allocated for each task. Using an adaptive vision system provides a more flexible solution as its analysis strategy can be changed according to the information available concerning the execution context. The authors describe how to create and evaluate a visual attention system tailored for interacting with a computer vision system so that it adapts its processing according to the interest (or salience) of each element of the scene. The authors propose a new set of constraints called 'PAIRED' to evaluate the adequacy of a model with respect to its different applications. The authors then justify why dynamical systems are a good choice for visual attention simulation, and we show that predator-prey models provide good properties for simulating the dynamic competition between different kinds of information. They present different results (cross-correlation, Kullback-Leibler divergence, normalized scanpath salience) that demonstrate that, in spite of being fast and highly configurable, their results are as plausible as existing models designed for high biological fidelity.

INTRODUCTION

While machine vision systems are becoming increasingly powerful, in most regards they are still far inferior to their biological counterparts. In human, the mechanisms of evolution have generated the visual attention system which selects the most important information in order to reduce both cognitive load and scene understanding ambiguity. Thus, studying the biological systems and applying the findings to the construction of computational vision models and artificial vision systems are a promising way of advancing the field of machine vision.

DOI: 10.4018/978-1-4666-3994-2.ch022

In the field of scene analysis for computer vision, a trade-off must be found between the quality of the results expected, and the amount of computer resources allocated for each task. It is usually a design time decision, implemented through the choice of pre-defined algorithms and parameters. However, this way of doing it limits the generality of the system. Using an adaptive vision system provides a more flexible solution as its analysis strategy can be changed according to the information available concerning the execution context. As a consequence, such a system requires some kind of guiding mechanism to explore the scene faster and more efficiently.

In this chapter, we propose a first step to building a bridge between computer vision algorithms and visual attention. In particular, we will describe how to create and evaluate a visual attention system tailored for interacting with a computer vision system so that it adapts its processing according to the interest (or salience) of each element of the scene.

Somewhere in between hierarchical salience based and competitive distributed models, we propose a hierarchical yet competitive model. Our original approach allows us to generate the evolution of attentional focus points without the need of either saliency map or explicit inhibition of return mechanism. This new real-time computational model is based on a dynamical system. The use of such a complex system is justified by an adjustable trade-off between nondeterministic attentional behavior and properties of stability, reproducibility and reactiveness.

In the first two sections, we start by giving a brief overview of the main theories and concepts of human visual attention and we provide the forces and weaknesses of state of the art attention models. This analysis is based on their potential of integration into adaptable computer vision system. We propose a new set of constraints called 'PAIRED' to evaluate the adequacy of a model with respect to its different applications. In a third section, we provide an in-depth description of our model and its implementation. We justify why dynamical systems are a good choice for visual attention simulation, and we show that predator-prey models provide good properties for simulating the dynamic competition between different kinds of information. This dynamical system is also used to generate a focus point at each time step of the simulation. In order to show that our model can be integrated in an adaptable computer vision system, we show that this architecture is fast and allows a flexible real time visual attention simulation. In particular, we present a feedback mechanism used to change the scene exploration behavior of the model. This mechanism can be used to maximize the scene coverage (explore each and every part) or maximize focalization on a particular salient area (tracking).

In a last section we present the evaluation results of our model. Since the model is highly configurable, its evaluation will cover not only its plausibility (compared to human eye fixations), but also the influence of each parameter on a set of properties (stability, reproducibility, scene exploration, dynamic behavior). But let's start by presenting attention.

THEORIES AND MECHANISMS OF BIOLOGICAL VISUAL ATTENTION

Why Visual Attention?

Does what we see really represent the world? Our perception of scenes seems accurate, continuous and consistent; however, the study of the various components of our visual system reveals a very different situation.

Eye and retina do not capture a perfect representation of the world. Actually, the distribution of different photoreceptor cells on the retina is not homogeneous: the center of it (the *fovea*) contains many more cone cells than the periphery. One major consequence is that representation of scene in human vision is more accurate in the center than in