Chapter 18
Biologically Inspired Components in Embedded Vision Systems

Li-Minn Ang
Charles Sturt University, Australia

Kah Phooi Seng
Charles Sturt University, Australia

Christopher Wing Hong Ngau
Intel, Malaysia

ABSTRACT

Biological vision components like visual attention (VA) algorithms aim to mimic the mechanism of the human vision system. Often VA algorithms are complex and require high computational and memory requirements to be realized. In biologically-inspired vision and embedded systems, the computational capacity and memory resources are of a primary concern. This paper presents a discussion for implementing VA algorithms in embedded vision systems in a resource constrained environment. The authors survey various types of VA algorithms and identify potential techniques which can be implemented in embedded vision systems. Then, they propose a low complexity and low memory VA model based on a well-established mainstream VA model. The proposed model addresses critical factors in terms of algorithm complexity, memory requirements, computational speed, and salience prediction performance to ensure the reliability of the VA in a resource constrained environment. Finally a custom softcore microprocessor-based hardware implementation on a Field-Programmable Gate Array (FPGA) is used to verify the implementation feasibility of the presented model.

INTRODUCTION

The human visual system (HVS) consisting of the eye and the brain, is a highly complex physiological system that is responsible for gathering visual information and translating them into useful responses. Furthermore, the human eye is capable of capturing massive amounts of visual information with each glance, probably exceeding ten megabits per second in continuous viewing (Davies, 2005). To promote efficient utilization of processing resources, the brain only selects relevant information to be processed at a time. In the early cognitive study by Broadbent (Broadbent, 1985), information selection by the brain is illustrated with a computational HVS where information in the form of stimuli is selected via selective filters for processing. This selection quality which is rather subjective in nature is called salience. Since the introduction of the computational HVS model by Broadbent, researchers in the field of psychophysics and vision have poured in much dedicated interest and research in understanding the mechanism of the HVS; which eventually led to the first neurally plausible visual attention system (Koch & Ullman, 1985). Research on the underlying mechanism of visual attention began since the 1890s where psychological studies on the nature of attentive process were carried out (James, 1890). From the early studies, visual attention was found leading to the actions of perceiving, conceiving, distinguishing, and remembering parts of a visual scene. For a visual scene seen by an observer, attention will cause the allocation of processing resources to salient areas rather than the entire scene as a strategy to work around the limited processing capacity of the human brain (Pashler, 1998). As a result, parts of the visual scene can be broken down to a series of localized and computationally less demanding information for rapid processing (Itti, Koch, & Neibur, 1998). The purpose of attention is to shorten the reaction time for the HVS to process a visual scene by first attending to salient areas which would bring about a visual awareness and understanding of the entire scene (Itti, 2005; Rapantzikos, Avrithis, & Kollias, 2007).

Computational visual attention (CVA) can be defined as algorithmic models that provide means of selecting parts of the visual input information for further higher vision processing or investigation based on the principles of the human selective attention (Frintrop, 2011). In contrast to theoretical biological and mathematical computer vision attention models, CVA models as an overlapping effort from both branches allow the validation of the modeled attention by comparing the output of the model to the results of vision experiments with visual stimuli such as images and video frames as inputs (Rothenstein & Tsotsos, 2011). Features present in a visual stimulus can be divided into two categories of high-level and low-level. High level features allow a semantic bridge between the extracted visual information and the interpretation of same information by the user (Zheng, Li, Si, Lin, & Zhang, 2006); resulting in the understanding of the entire visual scene (Nystrom & Holmqvist, 2008). These features exist in various subjective forms and are often defined by data or image representations (Li, Su, Xing, & Fei-Fei, 2010). While high-level features provide a more accurate and semantic understanding of the visual scene, extraction of these features are time consuming and involves large training databases (Zheng, Li, Si, Lin, & Zhang, 2006; Li, Su, Xing, & Fei-Fei, 2010). In contrast, low-level features are distinctive attributes that can be readily and easily extracted from objects or regions contained in a visual scene. Low-level features, unlike the high-level features, contribute minimally to the understanding of the visual scene. Although unable to provide a full semantic understanding of the entire visual scene, the low-level features play an important role in providing the quality of being unique for the information, locations, and objects contained in the visual stimuli (Wolfe, 1998). Color, edge, luminance, orientation, scale, shape, and texture are among the low-level features a human can efficiently detect with ease.

Related Content

Detecting Corners for 3D Objects
Misbah Irshad, Irfa Elahiand Malik Zawwar Hussain (2013). *Intelligent Computer Vision and Image Processing: Innovation, Application, and Design (pp. 272-280).*
www.irma-international.org/chapter/detecting-corners-objects/77046

"Scanning from Heating" and "Shape from Fluorescence": Two Non-Conventional Imaging Systems for 3D Digitization of Transparent Objects
Fabrice Mériaudeau, R. Rantoson, G. Eren, L. Sanchez-Sécades, O. Aubreton, A. Bajard, D. Fofi, I. Mohammed, O. Morel, C. Stolzand F. Truchetet (2012). *Depth Map and 3D Imaging Applications: Algorithms and Technologies (pp. 229-243).*
www.irma-international.org/chapter/scanning-heating-shape-fluorescence/60268

Ensemble Classification System for Scientific Chart Recognition from PDF Files
S. Nagarajanand V. Karthikeyani (2012). *International Journal of Computer Vision and Image Processing (pp. 1-10).*
www.irma-international.org/article/ensemble-classification-system-scientific-chart/75766

Entropy based Range Optimized Brightness Preserved Histogram-Equalization for Image Contrast Enhancement
Krishna Gopal Dhal, Sankhadip Sen, Kaustav Sarkarand Sanjoy Das (2016). *International Journal of Computer Vision and Image Processing (pp. 59-72).*
www.irma-international.org/article/entropy-based-range-optimized-brightness-preserved-histogram-equalization-for-image-contrast-enhancement/170596

Advances in Artificial Intelligence for Image Processing: Techniques, Applications, and Optimization
Sampath Boopathi, Binay Kumar Pandeyand Digvijay Pandey (2023). *Handbook of Research on Thrust Technologies' Effect on Image Processing (pp. 73-95).*
www.irma-international.org/chapter/advances-in-artificial-intelligence-for-image-processing/328027