ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

A Review of Modelling Visual Attention using Computational Cognitive Neuroscience for Machine Vision

Manjunath R Kounte[1], Dr. B.K.Sujatha[2]

Assistant Professor, Department of ECE, REVA I.T.M., Bangalore, India [1]

Professor, Department of TCE, MSRIT, Bangalore, India [2]

Abstract: In this review paper, we focus on learning how visual attention is beneficial to the machine vision community. Also we will explain why attention is considered a selective process. This review includes the survey of visual attention and its applications in various domains like computer vision, robotics, Advertising, Finding tumours in mammograms, Retinal prostheses and many more. This review highlights how in the last 25 years research on attention has characterized into multiple ways. This review also includes study of computational science, cognitive science and computational cognitive neuroscience with emphasis on modelling of visual attention by Psychophysics, Computational Methods and Neurophysiology. Comparison with other methods of computation modelling shows that Computational Cognitive Neuroscience approach in the best and most evolved system for exhibiting intelligence in real time scenarios. Finally survey of simulators such as GENESIS, NEURON, NEUROGRID, SPIN-MAKER, Neo-Cortical Simulator, Brain, Neural Simulation Tool and NeMo and CARLsim is done and we conclude that CARLsim is most suitable which allows for details and parameters to be specified at the synapse, neuron and network level.

Keywords: Visual Attention, Computational Sciences, Cognitive Sciences, Computational Cognitive Neuroscience, Psychophysics, Computational Modelling, Neurophysiology, CARLsim

I. INTRODUCTION

The human nervous system is a network of very large size and complexity, and extreme size and complexity and understanding the principles of brain processing by reverse engineering neural circuits and computational modelling is one of the biggest challenges of the Twenty-first century [1].

The purpose of this paper is to review the progress being done in the field of artificial intelligence where systems or machines learn on its own and imitate the human beings [2]. We understand the cognition behind attention and visual attention in particular and in this paper we have computationally modeled specific mechanisms of attention general and visual attention in particular. This paper shows that how visual attention is improved by providing proper attention to computational modelling. This work spans the fields of, Computational Sciences, Cognitive Sciences, Computational Cognitive Neuroscience, Psychophysics, Computational Modelling, and Neurophysiology.

Section II makes a basic study of visual information processing and three levels of interpretation which includes: Computational theory, Representations and algorithms and Hardware implementation.

The BDI (belief, Desire and Intention) Model is the basic model for any attention and perception. And it's important the accumulation of only necessary information to the machine should be identified or else the machine will be flooded with chunks of irrelevant information. Section III describes basic advantages and limitations of computational sciences and cognitive science. And Introduction to the world of computational cognitive neuroscience.
Section IV has a detailed comparison of Visual attention modelling using various possible methods which include Psychophysics, Computational Methods and Neurophysiology. We will focus on computational modelling which can be further be divided into filter models, neural models and computational cognitive neuroscience models.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Section V identifies Spiking Neural Network Simulator as a tool for Implementation. To articulate significant conclusions to a wider audience, paper concludes with open research issues and future scope.

II. VISUAL ATTENTION

As we open our eyes, we experience the world filled with enormous amount of visual information. The visual experience is joy and fun for our eyes but it's a signal processing nightmare for our brains – there is continuous flow of visual information which comes at a rapid rate to our retinas needs and this has to be processed to extract the small portions of information that are important for our actions[3].

A lot of the philosophy of how vision was believed to work at the time is summarized into notion of the three levels of description [4] of a visual information processing system. These three levels, can be interpreted as:

• Computational theory: What is the goal of the computation (task) and what are the constraints that are known or can be brought to bear on the problem?

• Representations and algorithms: How are the input, output, and intermediate information represented and which algorithms are used to calculate the desired result?

• Hardware implementation: How are the representations and algorithms mapped onto actual hardware, e.g., a biological vision system or a specialized piece of silicon? Conversely, how can hardware constraints be used to guide the choice of representation and algorithm? With the increasing use of graphics chips (GPUs) and many-core architectures for computer vision, this question is again becoming quite relevant.

A qualitative definition of visual attention [5] is given as, taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others...

In [6], the scientist Von Helmholtz posited visual attention as an essential mechanism of visual perception and was mainly concerned with eye movements to spatial locations, or the "where" of visual attention. In essence, although visual attention can be consciously directed to peripheral objects, eye movement's reflect the will to inspect these objects in fine detail. In this sense, eye movements provide evidence of overt visual attention.

James defined attention mainly in terms of the "what", or the identity, meaning, or expectation associated with the focus of attention. James favoured the active and voluntary aspects of attention although he also recognized its passive, reflexive, non-voluntary and effortless qualities.

Both views of attention, which are not mutually exclusive, bear significantly on contemporary concepts of visual attention. The "what" and "where" of attention roughly correspond to foveal (James) and parafoveal (Von Helmholtz) aspects of visual attention, respectively

Selective visual attention provides the brain with a mechanism of focusing computational resources on one object at a time, either driven by low-level image properties (bottom-up attention) or based on a specific task (top-down attention). Moving the focus of attention to locations one by one enables sequential recognition of objects at these locations

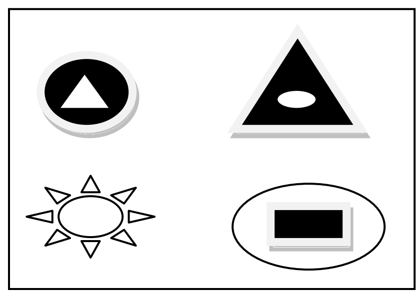

Fig 1: Vision definition

A) Vision before attention - Looking at Figure 1, we see something before we knew if that something included a black triangle surrounding a white Circle. Here the visual information is available "preattentively", before attention is directed to a locus or an object [7].

B) Vision with attention - Most of vision research generally take into consideration vision with attention and the subjects are generally asked to perform some task at the same time attending to an example of visual stimulus. Here we focus on altering a preattentive visual representation?

C) Vision after attention – Here the focus is on attention which is restricted in space and time, does it leave any presence, once it has been deployed away from a stimulus?

D) Vision without attention - In some situations it is meaningful to ask about the fate of stimuli that are never attended. This is related to, but not identical to, the question of vision before attention

III. COMPUTATIONAL COGNITIVE NEUROSCIENCE

A. Computational Sciences

Computational Science [8] is increasingly pushing the limits of what is possible in computing, and in some areas is

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

leading computational advances. Theoretical chemistry, Climate research, earth system research, computational fluid dynamics, genomics, astrophysics, proteomics, nanostructure physics, and high-energy physics are all examples. New Scientific computing platforms and high end infrastructures are making possible new kinds of experiments that would have been impossible to conduct only 10 years ago, changing the way scientists do science.

Our focus here is to discuss about the limitations of Artificial Intelligence. Ever since it was introduced, Artificial intelligence (AI) has been one of the most interesting and controversial domains of inquiry in computational science. Artificial Intelligence is defined as part of computational science concerned with designing robots and systems that exhibit the characteristics associated with human intelligence— understanding language, learning, analysing, reasoning, solving problems etc. Because of various intellectual challenges, and scope, the field has attracted researchers to take up research topics in this field.

There were two main approaches to AI

A Classical approach, which is based on mathematical approach in which ideas and concepts are represented by symbols such as words, phrases or sentences, which are then processed according to the rules of logic.

A Connectionist approach, based on artificial neural networks, which imitate or behave the way neurons work, and genetic algorithms, which imitate inheritance and fitness to evolve to much improved solutions to a problem with every next generation.

The field has had fair share of controversies also because of its social, ethical, and philosophical implications. Such controversies has affected the funding environment for AI and the objectives of many research programs.

We would like to discuss about the contribution of AI to develop human level intelligence. The issue is understanding how unintelligent components can combine to generate human-level intelligence *the intelligence problem*; the endeavour to understand how the human brain embodies a solution to this problem *understanding human intelligence*; and the project of making computers with human-level intelligence *human-level artificial intelligence*.

The issue is AI [9] should help develop a system exhibits human-level intelligence, i.e. it can deal with the same set of situations that a human can with the same level of competence. For example, the machines should have same kinds of conversationalist techniques as a typical human being. Situations which require physical manifestation should also considered to develop a system. It's not just challenges related to programming, without hand gestures and facial expressions, it would be difficult for a machine to converse as well as a human in many situations because hand gestures and facial expressions are part of many conversations for human beings. In the long term, it's very important that numerous possibilities of situations are taken into consideration to develop the systems. However, the current abilities of artificial systems are so far away from human-level that resolving this issue can generally be postponed for some time. One point that does follow from these reflections, though, is the inadequacy of the Turing Test. Just as the invention of the airplane was an advance in artificial flight without convincing a single person that it was a bird, it is often irrelevant whether a major step-step towards human-intelligence cons observers into believing a computer is a human[10].

B. Cognitive Sciences

The term Cognitive Science has defined as the study of Cognitive structures and processes in mind or brain, [11].

Cognitive Science is an interdisciplinary field which includes a wide array of fields. These include philosophy, psychology, linguistics, artificial intelligence, robotics, and neuroscience. The Common interest each field brings with it a unique set of tools and perspectives that has emerged by the intersection of existing branches of science such as Psychology, Linguistics, Computer Science, Philosophy and Physiology. "The shared interest that has produced this coalition understands the nature of the mind", [12].

The term cognitive science refers not so much to the sum of all these disciplines but to their intersection or converging work on specific problems. In this sense, cognitive science is not a unified field of study like each of the disciplines themselves, but a collaborative effort among researchers working in the various fields. The glue that holds cognitive science together is the topic of mind and, for the most part, the use of scientific methods

Origin of Cognitive Science began in mid-1950 when the researches began to treat mind not as a machine but as a biological system which has several non-linear responses and can be analyzed based on complex representations and computational

The moral in so far discussion is the human mind can be investigated based on computations. Hence Computer Science in general and Artificial Intelligence in particular have come to play a central role in Cognitive Science.

C. Computational Cognitive Neuroscience

Cognitive Computing is an advanced topic that goes beyond AI and human-computer interaction by combining the concepts of perception, memory, attention, language intelligence and consciousness, [13].

Cognitive Computing is a study of top-down, global, unifying theories that explain observed cognitive phenomena ("mind"), that are consistent with known bottom-up neurobiological facts (the "brain"), that are computationally feasible, and that are mathematically principled. Cognitive Computing is a search for computer science-type software/hardware elements that are consistent with known neurobiological facts about the brain and give rise to

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

observed mental processes of perception, memory, language, intelligence, and, eventually, consciousness. Very simply speaking, "Cognitive Computing is when computer science meets neuroscience to explain and implement Psychology, [14]". Both AI and *Neural networks* (NN) fail to replicate the phenomenon of thinking, though they have considered brain as the starting point to explain Cognitive phenomena. In contrast Cognitive Computing is about learning how brain operates, designing algorithms, reverse engineering and testing plausible models. "Cognitive Computing is about engineering the mind by reverse engineering the brain, [14]". The principles and methodologies of Cognitive sciences are being implemented to robots, so as to create humanoid robots. These robots are not just software agents but are mobile robots and autonomous vehicles.

In Neurolinguistics, Cognitive Neuroscience is used to solve neurocognitive disorders and help in computational modeling. There are many projects using Cognitive computing for investigation of different languages like Spanish, Russian, Japanese, Portuguese, Dutch etc., [15].

Computational Cognitive Neuroscience is also used in investigation of hearing and speech. Brain imaging technique also uses Cognitive computing to see how a man speaks in pathological conditions as compared to healthy conditions.

IV. MODELLING OF VISUAL ATTENTION

The advantages of Modelling Visual Attention includes in the field of Computer Vision and Graphics like Image segmentation, Image quality assessment, Image matching, Image rendering, Image and video compression, Image thumb nailing, Image super-resolution , Image re-targeting (thumb nailing), Image super resolution, Video summarization, Scene classification, Object detection, Salient object detection, Object recognition, Visual tracking, Dynamic lighting, Video shot detection, Interest point detection, Automatic collage creation Face segmentation and tracking[16].

In the field of Robotics like Active vision, Robot Localization, Robot Navigation, Human-robot interaction, Synthetic vision for simulated actors. In various other fields like Advertising, Finding tumours in mammograms, Retinal prostheses etc.

In human beings, the attention is facilitated by a retina that has evolved a high-resolution central fovea and a low-resolution periphery. The important parts of scene gathering and collection of important information is guided by the anatomical structure of the retina. We focus on the Computational Modelling of this interesting field.

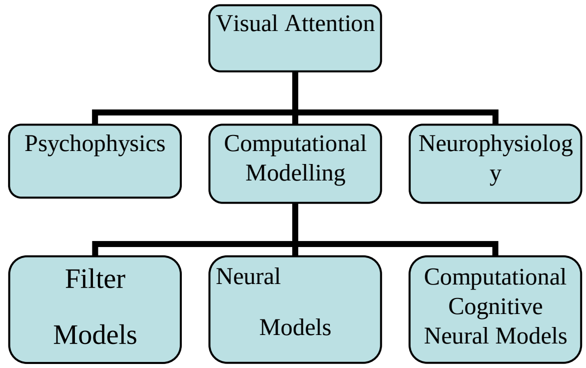

Fig 2: Taxonomy of Visual Attention Models

As Shown in the figure 2, there are various approaches to the modelling of the visual attention. Research is being carried out by scientists from numerous domain of science. Foremost among them are the Psychologists who have studied behavioural correlates of visual attention such as change blindness [18] [19], inattentional blindness [20], and attentional blink [21]. Whereas Neurophysiologists have shown that how the neurons accommodate themselves to better represent objects of interest [24] [25].

In Computational Modelling approach, the filter models have built models that can compute saliency maps and realize the Visual attention based on Top-Down Attentional Models and Bottom-Up Models. In neural network models, the approach is to simulate and explain attentional behaviours (e.g., [22] [23]).

Although there are many models available now in the research areas mentioned above, here we propose a new model based on computational Cognitive Neuroscience to build systems capable of working in real-time.

V. SIMULATION OF VISUAL ATTENTION

Despite funding from various agencies and increase in computer power, constructing a model the size of a human-brain will require a large scale increase in magnitude in memory capacity, processing capacity and increase in speed of communication[26]. Conventional computer architecture like von- Neumann or Princeton architecture may not be the appropriate architecture for modelling a brain. Unlike a conventional computer, the brain is a massively parallel processing, analog, fault-tolerant, and selective system that does not rely on programmed instructions [27].

Hence there is a need within the computational cognitive neuroscience community for abstractions and simulation environments that support modelling at a large-scale. In particular, we consider large-scale network models of spiking neurons, which demonstrate important properties of neurobiological information processing, such as temporal dynamics, precise timing, and brain rhythms [28]. There are many spiking simulators, which are currently available, which fall into different categories based on abstraction level and the hardware used to develop the computer. Simulators, such as GENESIS and NEURON, incorporate molecular, models of axons and dendrites from anatomical observations, a major goal of these models is to study detailed ionic channels and their influence on neuronal firing behaviour. The problem with these simulators is high computation costs. Here these components are modelled with differential equations having time steps less than 1ms. Hence, large-scale simulation of the brain is extremely challenging at this level.

Simulators like NEUROGRID, and SPIN- NAKER, are efficient enough to run large-scale networks of spiking neurons, but they require specialized hardware Therefore, these systems are not readily available to the computational cognitive neuroscience community [29]. Simulation environments, such as the neo cortical simulator, Brian, Neural Simulation Tool, NeMo, are specifically designed for developing spiking neuron networks. However, each simulator environment has different trade-offs in speed, realism, flexibility, maximum network size, etc. Our approach is to use a simulator that is easy to use and yet provide significant computational performance. **CARLsim** is an efficient C/C++-based Spiking Neural Network (SNN) simulator that allows execution on both generic x86 CPUs and standard off-the-shelf GPUs. The simulator provides a PyNN-like programming interface, which allows for details and parameters to be specified at the synapse, neuron, and network level.

VI. CONCLUSION

In this paper we have given a Critical review of Visual Attention Modelling using different streams arising because of Psychophysics, Computational Methods like Filter Models & Neural Models, and Neurophysiology. We highlighted the emergence of Cognitive sciences, computational Intelligence, and Neuroscience, also the emergence of new field of science which combines psychology, computer science, mathematics and neuroscience and Electronic Hardware Design. The cognitive neuroscience computing is at a verge to give rise for many potential mainstream applications.

The review also focused on modelling of visual attention using computational cognitive neuroscience. Survey of Spiking neural network simulators showed that CARLsim Simulator is the best option for implementation of Visual Attention using Computational Cognitive Neuroscience for real time scenarios.

In future, we intend to simulate large scale cortical networks and find efficient visual attention principles to apply for real life applications and scenarios using computational cognitive neuroscience principles.

REFERENCES

1) Micah Richert, Jayram Moorkanikara, Nageswaran, Nikil Dutt, and Jeffrey L. Krichmar, "An efficient simulation environment for modeling large-scale cortical processing" Frontiers in Neuroinformatics, Sep 2011
2) Boden, M. A. Mind As Machine: A History of Cognitive Science. Oxford University Press, Oxford, England. 2006
3) Dirk Walther "Interactions of Visual Attention and Object Recognition: Computational Modeling, Algorithms, and Psychophysics" California Institute of Technology, Rep. Mar 7, 2007
4) Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W. H. Freeman, San Francisco.
5) *The Works of William James*, Cambridge, MA and London: Harvard University Press, 17 vol., 1975
6) Hermann von Helmholtz and James P.C. Southall, "Helmholtz's Treatise on Physiological Optics" 1975
7) Wolfe J. "Visual attention". 2nd ed. San Diego, CA: Academic Press; 2000. p. 335-386
8) "Towards 2020 Science", Microsoft research Foundation.
9) Committee on Innovations in Computing and Communications: Lessons from History, National Research Council, "Funding a Revolution: Government Support for Computing Research" National Academy of Press. 1999
10) Nicholas L. Cassimatis, "Artificial Intelligence and Cognitive Modelling Have the Same Problem" , Theoretical Foundations of Artificial General Intelligence Atlantis Thinking Machines Volume 4, 2012, pp. 11-24
11) Daniel Terdiman, "IBM: Computing rivalling human brain may be ready by 2019" on http://news.cnet.com/8301-13772_3-10400362-52.html
12) David Templeton, "CMU finds human brains similarly organized", Pittsburgh Post-Gazette (January 4, 2008).
13) O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., and Contributors (2012). *Computational Cognitive Neuroscience*. Wiki Book, 1st Edition
14) Cognitive Computing 2007: May 2-3, 2007, UC Berkeley CITRIS and NERSC Sponsored Research
15) Speech and Language : language group in NACS,http://www.nacs.umd.edu/res/speech.html
16) Ali Borji, Laurent Itti,," State-of-the-art in Visual Attention Modelling" IEEE Transactions On Pattern Analysis And Machine Intelligence, 2010
17) Richard Szeliski "Computer Vision: Algorithms and Applications " Springer Publication ,2010
18) D.J. Simons and D.T. Levin, "Failure to Detect Changes to Attended Objects," Investigative Ophthalmology & Visual Science, vol. 38, no. 4, pp. 3273, 1997.
19) R. A. Rensink, "How Much of a Scene is Seen - the Role of Attention in Scene Perception", Investigative Ophthalmology & Visual Science, vol. 38, 1997.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

20) D.J. Simons and C.F. Chabris, "Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events," Perception, vol. 28, no. 9, pp. 1059-1074, 1999.
21) J.E. Raymond, K.L. Shapiro, and K.M. Arnell, "Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink?" Journal of exp. psych., vol. 18, no 3, pp. 849-60, 1992.
22) L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Transactions on PAMI, vol. 20, no. 11, pp. 1254-1259, 1998.
23) J.K. Tsotsos, S.M. Culhane, W.Y.K. Wai, Y. Lai, N. Davis, and F. Nuflo, "Modeling Visual Attention via Selective Tuning," Artif. Intell., vol. 78, no. 1-2, pp. 507-545, 1995.
24) R.D. Rimey and C.M. Brown, "Controlling Eye Movements with Hidden Markov Models," IJCV, vol. 7, no. 1, pp. 47-65, 1991.
25) S. Treue, "Neural Correlates of Attention in Primate Visual Cortex," Trends in Neurosciences, vol. 24, no. 5, pp. 295-300, 2001.
26) Jeffrey L. Krichmar, Nikil Dutt, Jayram M. Nageswaran, Micah Richert, "Neuromorphic Modeling Abstractions and Simulation of Large-Scale Cortical Networks" IEEE 2011
27) Izhikevich, E. M.. "Simple model of spiking neurons." Ieee Transactions on Neural Networks no. 14 (6):1569-1572. doi: Doi 10.1109/Tnn.2003.820440. 2003
28) Izhikevich, E. M.. "Which model to use for cortical spiking neurons?" Ieee Transactions on Neural Networks no. 15 (5):1063-1070. doi: Doi 10.1109/Tnn.2004.832719 2004
29) Nageswaran, J. M., N. Dutt, J. L. Krichmar, A. Nicolau, and A. V. Veidenbaum. 2009. "A configurable simulation environment for the efficient simulation of large-scale spiking neural networks on graphics processors." Neural Networks no. 22 (5-6):791-800. doi: DOI 10.1016/j.neunet.2009.06.028
30) Richert, M., J. M. Nageswaran, N. Dutt, and J. L. Krichmar. 2011. "An efficient simulation environment for modeling large-scale cortical processing." Front Neuroinform no. 5:19. doi: 10.3389/fninf.2011.00019