

Simulating and Analyzing IF and LIF Neuron Models for Selective Visual Attention

Anu Yadav

Department of Computer Science & Engineering
KIET Group of Institutions
Ghaziabad, India

Sanjeev Kumar Yadav

Department of Computer Science & Engineering
KIET Group of Institutions
Ghaziabad, India

ABSTRACT

By nature biological systems have the capacity to selectively guide their attention on some portion of visual scene and process only that selected information for further processing. The similar abilities are also required in artificial vision systems so that they can become more efficient and realistic as biological systems and in future manual supervisory and detection work can be replaced by artificial visual systems. In this paper concepts of visual information processing and selective visual attention are implemented and analyzed. The results from two neuron models viz. IF and LIF models are compared and an analysis is made. In visual selective attention model the attention is guided towards only the most salient object while ignoring the less salient objects this way the time taken to process the image is reduced as large amount of information which is not much relevant is rejected for further processing. The study made in this paper will be proving to very effective as it will enhance the reliability of the system. The visual system can be used in robot vision system, can be deployed in choke points of city and major functions to detect accidents etc. This concept is also useful in marketing to make a few second advertisements more effective.

Keywords

IF-integrate and fire model, LIF-leaky integrate and fire model, Selective visual attention, salient objects, robot vision system, supervisory and detection work

1. INTRODUCTION

Visual attention is naturally adapted by biological systems to deal with large amount of input visual information. It is a cognitive process which includes selectively concentrating on salient objects and ignoring less important parts i.e. filtering out irrelevant information of image and thus helps in reduction of large amount of visual information given as input [9]. It is a cognitive process, helps in making visual system more efficient by reducing the amount of information needed to processes so that limited computational resources can be provided with most important visual information [4].

According Tsotsos et al. [11], visual processes selects a part of visual information received as input and uses only these selected particular pieces of information in further processing. Due to limited capacity of processing, all objects in visual space compete for neural representation. This competition can be biased by both bottom up and top down approach of selective attention [6].

There are many applications of selective visual attention. It can be used in machine vision, especially in a robot vision system for target selection [11], in quickly detect prey and predators in visual world [5], in biologists study on patients of amnesia to analyze lobes of brain which causes attention disorder [13], in

marketing to make an advertisement poster and few second advertisements more attractive and in selecting specific colour, font and layout which catch more eyes, it can be deployed in major choke points of city and major functions to detect accidents etc., and in highlighting the players and suppressing background regions in playground.

2. VISUAL SELECTIVE ATTENTION

It is a process which helps human to work more efficiently by directing the attention to region of interest in visual environment and providing only the most important information to limited computational for further processing[4].

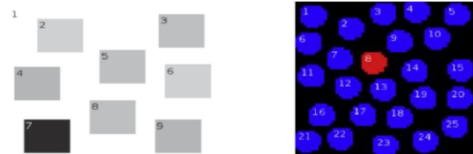


Fig 1: Input images showing different patterns [9].

In Fig 1(a), naturally attention is directed towards black box, object number 7. In Fig 1(b), attention is directed towards red ball, object number 8. The attention is directed towards most contrasting object.

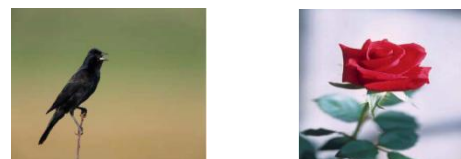


Fig 2: Input images of real life objects [9].

In Fig 2(a), attention is directed towards object black crow. In Fig 2(b), attention is directed towards object red rose. It shows the concept of visual selective attention is also applicable for real life situations.

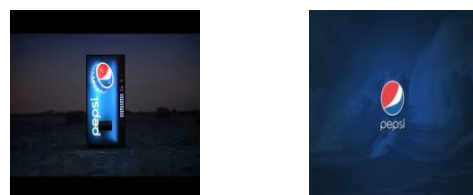


Fig 3: Input Images Showing Advertisement Posters.

In Fig 3(a) naturally attention is directed towards the object cold drink bottle. In Fig 3(b), attention is directed towards icon showing the brand name. It shows that the advertisement posters are build considering the contrast of the object with the background so that the attention is delivered quickly.

Visual selective attention can be mainly divided into two parts:

- *Top down attention /task dependent attention*

This kind of attention is guided by combined information from retina and visual cortical areas as well as feedback signals from area outside the visual cortex [9]. It depends on the task which is in hand. Eg. Search for black T-Shirt man in the garden [5].

It depends on factors, like task demands, mental state and personal experience [4] and hence takes more amount of time as compared to bottom-up attention [5]. It mainly functions in response to internal goals, that are mostly originates from parietal and frontal lobes of the brain [7].

- *Bottom-up attention/ scene dependent attention*

In this, simple features such as intensity, motion, color, orientation and others can be extracted from input image [9] and on the basis of these features visual attention is guided. It is faster and more automatic process as compared to top-down attention [7].

It is less associates with higher areas of brain so easier to stimulate. The attention exclusively depends on external stimuli and is not guided by any task related influence. Eg. A red flower in a green tree will draw attention even if we are searching for a butterfly.

Both mechanisms of attention top down and bottom up can operate in parallel [5].

There are three main approaches used for machine visual attention models:

A. *Location based/spatial attention*

Generally implemented by using WTA (winner takes all) mechanism. In this only a single neuron wins and only that neuron is activated or only a small area around that neuron is considered as a focus of attention, but not whole object or component [9].

B. *Object based attention*

It considers object as a basic unit and whole object competes for attention and visual attention is delivered to salient object or component. For using these models, the model must have an embedded segmentation function [9].

C. *Feature based attention (FBA)*

It is deployed to specific aspects (color, contrast, orientation) of visual inputs, and not considers their actual location. We use FBA in case, where we know a defining feature of an object, and do not know exact location of that object. E.g. The text book is black color and it is somewhere in bookshelf [3].

By focusing on particular kinds of information and attention we can optimize our search process. Spatial attention is concerned with a particular location and guides us to focus on that particular location, FBA directs us to focus on particular features in the visual input [3]. E.g. waiting for a friend in library, we may deploy spatial attention to the door and deploy feature based attention to red objects, if he has worn red shirts [3].

3. MODELING USING INTERGRATE & FIRE NEURON MODEL

Quiles et al. [9] proposed an object based visual selection model. It suggested that objects can be represented by temporal correlation and synchronization of the firing activities of neurons, which codes different features of an object. It is a oscillatory correlation model built on a network of integrate and fire neurons for visual selection. In oscillatory correlation model

neurons representing different features of same object are synchronized and neurons representing different objects are desynchronized. Synchronization works for grouping neurons into objects while desynchronization distinguishes one group of synchronized neurons from another.

As system runs, each group of neurons representing same object is synchronized due to cooperative mechanism among neighbor neurons. At same time competitive mechanism among neurons is introduced by long range connection among neurons, by this competitive mechanism firing frequencies of those neurons which represent salient object is increased and firing frequencies of neurons which represent background is decreased. The firing frequencies of neurons are controlled by external input and coupling term.

This model, considers visual attributes such as intensity and contrast of colors for visual selection. This model encodes contrast of features rather than their absolute level and this result in easier search when target-distracter difference is large. The output is not a single winning neuron but a group of winning neurons shows a salient object.

The model also doesn't require any global coordinating neuron, as it uses both excitatory and inhibitory connections. The excitatory connection tends to synchronize neurons which represent same object and inhibitory connections inhibit neurons representing non salient objects. In this model, the neurons tends to self organize on the basis of synchronization. The coupling weights between neurons are determined by external stimuli. The given stimulus evokes process of oscillation and synchronization among neurons, representing same object.

The IF neuron model is one of the simplest models of neural dynamics and defined by equation [9]:

$$\frac{dV}{dt} = C \quad (1)$$

Where,

'V' represents phase/ potential of the neuron. C represents rate of growth of phase/potential of neuron.

In this model, the neuron is characterized by a single variable, which is membrane potential. When potential of this neuron reaches a certain threshold value, the neuron fires and its potential is reset to a zero value. The phase change of each neuron, in a network of IF neurons, is given by

$$\frac{dV_i}{dt} = C + \Delta V_i \quad (2)$$

Where,

'i' represents oscillator index &

' ΔV_i ' is a coupling term.

If a network consists of only two neurons then $i = \{1, 2\}$.

Now, when phase of oscillator O_i reaches a certain threshold value ($V_i = 1$) then its phase and phase of all other neurons which are connected to it will be changed by following equation:-

$$V_i \geq 1 \Rightarrow \begin{cases} V_i = 0 \\ V_j = V_j + \Delta(V_j) \end{cases} \quad (3)$$

Where,

V_j , are phases of oscillators O_j which are connected to the firing oscillator O_j .

$\Delta(V_j)$, is phase change of oscillator O_j due to pulse transferred by O_i to O_j & is given by:

$$\Delta(V) = \varepsilon V \quad (4)$$

Where

' ε ' is a constant which defines the excitatory coupling strength between i & j neurons.

In a network of neurons when any neuron fires it sends a pulse to all other neurons which are connected to it. According to the types of synapses, the pulse transferred can be categorized as excitatory (positive) and inhibitory (negative). If it increases the potential of the neuron which receives it and thus helps in firing of that neuron then the pulse is excitatory.

If the pulse is inhibitory, it inhibits and delays the next firing of neuron which receives it. Excitatory coupling helps in synchronizing and tend to synchronize the group of connected neurons. On the other hand, inhibitory synapses prevent synchronization.

All neurons in a network of IF neurons evolve linearly, the case of nonlinearity comes only when any neuron fires. Hence continuous model can be transformed to discrete one by considering only instances when any neuron fires.

Quiles et al. [9] proved that, if IF neurons are connected with excitatory connections, no matter how weak the coupling strength is, it will be definitely synchronized as time tends to infinity. A model which is formed by a two dimensional network of IF neurons with only two types of connections: excitatory short range & inhibitory long range. The network is defined by –

$$\frac{dV_i}{dt} = C_i(t) + E_i(t) - I_i(t) \quad (5)$$

Where

' i ' - represents neuron index. However in a 2 D network, each neuron has two indices: line & column. But for simplicity, here only one index for each neuron is used.

V_i - Neuron's phase.

$C_i(t)$ - denotes phase growing velocity of neuron, which controls the firing frequency of neuron ' i '.

$E_i(t)$ - denotes the excitatory coupling.

$I_i(t)$ - defines the inhibitory coupling among neurons.

The excitatory coupling $E_i(t)$ is defined as[9]-

$$E_i(t) = \varepsilon V_i(t) \sum_{j \in \Delta_i} \delta(t - t_j) \quad (6)$$

Where

δ - is Dirac delta function.

t_j - is the instant when neuron j fires.

Δ_i - defines the excitatory cooperation neighborhood of neuron i .

In more specific terms, any neuron ' i ' establishes an excitatory coupling to any other neuron j , only if j is one among the eight nearest neighbors of neuron i and the feature difference between these two neurons is below certain threshold value. ε - is a constant, which represents the excitatory coupling strength between two neurons i & j .

The inhibitory coupling $I_i(t)$ is defined as[9]-

$$I_i(t) = V_i(t) \sum_{j \in \Lambda_i} \sigma_{ij} \delta(t - t_j) \quad (7)$$

Where

Λ_i - represents the competitive neighborhood of neuron i .

σ_{ij} - represents inhibitory coupling strength between two neurons i & j , which is given by following equation[9]:

$$\sigma_{ij} = C_Y \exp\left(-\sum_k C_{ij}^k F_{ij}^k\right) \quad (8)$$

Where

C_Y is a constant and $C_Y \in [0, 1]$.

F_{ij}^k & C_{ij}^k represents the contrast of feature k and its weight respectively.

$k \in [1, n]$, denotes the feature (colour, intensity, orientation etc.).

The contrast F_{ij}^k can be defined as the absolute difference of the feature k between neurons i & j i.e.

$$F_{ij}^k = |F_i^k - F_j^k| \quad (9)$$

For object selection task, we need a mechanism which controls the firing frequency of each neuron. Firing frequency of the neuron is increased if that neuron is the part of salient object, otherwise firing frequency is decreased.

Firing frequency can be controlled by modeling the $C_i(t)$ term. Every time, when a neuron i fires, it increases its own $C_i(t)$ value by[9]-

$$C_i(t) = C_i(t - 1) + \Omega(C_{max} - C_i(t - 1)) \quad (10)$$

Where

Ω is a constant, represents the potential gain factor (phase shift).

C_{max} is a constant, represents maximum value that can be holded by $C_i(t)$.

Value of $C_j(t)$ can be decreased by the following equation[9]-

$$C_j(t) = C_j(t - 1) + \varepsilon \sigma_{ij} (C_{min} - C_j(t - 1)) \quad (11)$$

Where

C_{min} - is the minimum value that can be holded by $C_j(t)$.

When an inhibitory signal is received by any neuron j , from any other neuron i , the value of its $C_j(t)$ can be decreased according to above equation.

4. MODELING USING LEAKY INTEGRATE & FIRE MODELS

It was one of the earliest models of the neuron. In this the neuron was represented with respect to time by

$$I(t) = C \frac{d\phi}{dt} \quad (12)$$

When any stimulus (input current) is presented, as a result of it the membrane potential of the neuron increases with respect to time until it reaches a certain threshold value, after reaching threshold value a spike is generated, the neuron fires and transfers some chemicals to other neurons which are connected to it and then potential of the neuron is reset to rest potential.

This model also accounts for the memory problem which was not considered in the Integrate and fire neuron model. In this the problem is solved by introducing a ‘leak’ term to membrane potential. This leak represents the amount of ions that are diffused through membrane when threshold is not reached. The neuron is represented as-

$$I(t) - \frac{V(t)}{R} = C \frac{dV(t)}{dt} \quad (13)$$

Where ‘ R ’ represents membrane resistance as the membrane is not a perfect insulator, however it is assumed as a perfect insulator in Integrate and fire neuron model. If the input current exceeds a threshold value, it fires otherwise it leaks any change in potential value.

5. COMPUTER BASED SIMULATIONS

Take an input image, construct a network such that each neuron in the network represents a pixel of the input image i.e. each pixel in the input image, stimulates the corresponding neuron in the network. The connections among neurons in the network are set by this stimuli information. The excitatory connection between any pair of nearest neighbor neurons is turned on if both of those neurons capture the similar input features i.e. feature contrast is below a certain threshold value; otherwise, the connection between neurons is cut.

These cooperative connections tend to synchronize each group of connected neurons. When the network is stimulated with input image, the long range inhibitory connections are determined on the basis of the contrast of the input features. If any two neurons are less contrasting then the value of feature contrast will be small however the inhibition between these neurons will be strong. If the feature difference between two neurons is large, then the inhibitory signal between them will be of weak strength.

With this procedure, the excitatory coupling synchronize the group of coherent neurons and thus segment the input image into different objects, on the other hand the inhibitory connections decrease the firing rate of neurons which have lower color and intensity contrast to other parts of the image.

As a result of this cooperative and competitive process, the salient object gets highlighted and the non salient objects become silent.

5.1 Using Integrate & Fire Model

Extract the R , G , B components of the coloured input image. Intensity can be calculated as average of all three color components for each neuron. Normalise these features between 0 and 1. The corresponding features of each pixel will stimulate respective neurons of the network. The values of the parameters taken in the simulation results are as follows: $C_{min} = 0.1$; $C_{max} = 2.0$; $\epsilon = 0.2$; $C_y = 0.3$; $R = 10e1$;

$C_0 = 1.1$; $C_{ij} = 1$ for all k .

The results of implementation of visual selective attention using LIF model are presented. Different images are taken and using the built model the firing frequency and the potential of each neuron with respect to time are analyzed.



Fig 4: Input image and input image showing pixel locations for which potential values are recorded.

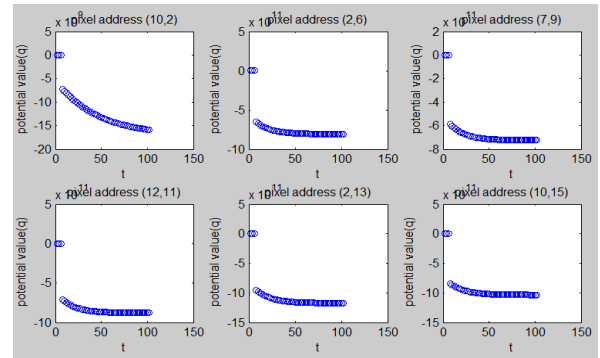


Fig 5: Potential Values of Pixels which were selected with respect to time.

From above results it is clear that object corresponding to pixel(7,9) i.e. blue object is salient as we can see that its potential is decreasing least and it will fire in next coming cycles of ‘ t ’ (time).

5.2 Comparison Using IF and LIF Model



Fig 6: Input image and input image showing pixel locations for which potential values are recorded.

5.2.1 Result by IF model

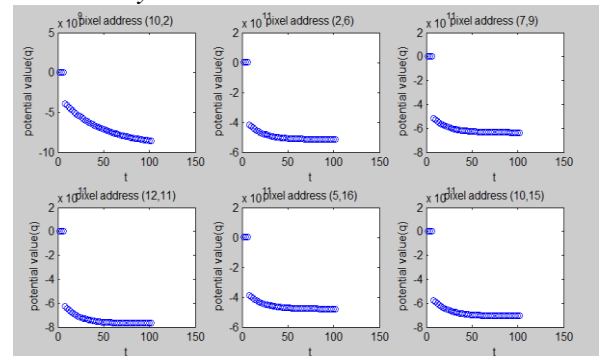


Fig 7: Potential Values of Pixels which were selected with respect to time.

5.2.2 Result by LIF model

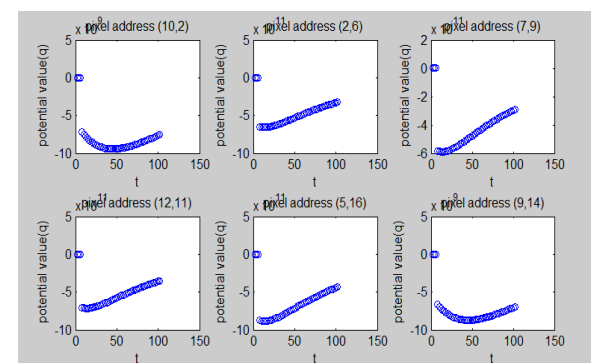


Fig 8: Potential Values of Pixels which were selected with respect to time.

From Fig 7 and Fig 8 it can be clearly seen that the LIF model is converging fast, as the potential value of neurons is growing fast and approaching to threshold frequency again after firing with a

great speed. So it will converge within less number of cycles of t' i.e. within less amount of time. As it can quickly detect the salient object and also accurately as proved in above example and it is more realistic (amount of time taken is more similar to biological systems as compared to integrate and fire model) and also produces similar results to biological systems.

6. CONCLUSION AND FUTURE SCOPE

The models are implemented and observed that the LIF model converges faster as compared to IF model so produces results in lesser time closely similar to biological systems as compared with IF model. However the results are subjective as high processing power was required to implement the model at large since as it is a one pixel- one neuron model, so number of neurons in the network should be equal to the number of pixels in the image.

Some limitations are observed during implementation of the model:

Here we simulate for on a single layer processing of input while in biological network the computation is done on many layers as each neuron can receive input from any or all pixels of image.

In future the visual selective attention can be applied to a video or any dynamic scene. The above mentioned limitations can also be removed to get more realistic results and reliability of the system can be enhanced.

7. REFERENCES

- [1] Alwyn Scott, *Neuroscience, A Mathematical Primer*, Springer- Verlag New York, 2002.
- [2] Becker S., *Modelling the Mind: From Circuits to Systems*, Chapter 2 in *New Directions in Statistical Signal Processing: From Systems to brain*, Haykin S. et.al. MIT Press, 2005.
- [3] Carrasco M, "Visual attention: The past 25 years, vision research 51", Elsevier Ltd.,2011, pp. 1484-1525.
- [4] Chen D and Chu H, "Scale invariant amplitude spectrum modulation for visual saliency detection", *IEEE Transactions on Neural Networks and Learning Systems*, Vol 23, No. 8, 2012, pp. 1206-1214.
- [5] Itti L and Koch C, "Computational Modeling of Visual Attention", *Macmillan Magazines Ltd.*, Vol. 2, 2001, pp. 194-203.
- [6] Kastner S and Ungerleider L G, "Mechanisms of visual attention in the human cortex", *Annual Review of Neuroscience*, Vol 23, 2000, pp. 315-341.
- [7] Neokleous K C and Schizas C N, "Computational modeling of visual selective attention", In *Proceedings of the European Future Technologies Conference and Exhibition*, 2011, pp. 244-245.
- [8] Niebur E. and Koch C., "A Model for the neuronal implementation of selective visual attention based on temporal correlation among neurons", *Journal of Computational Neuroscience* 1, pp. 141-158, 1994.
- [9] Quiles M. G., Zhao L., Breve F.A. and Romero R. A. F., "A network of integrate and fire neurons for visual selection", *Elsevier journal, Neurocomputing* 72, 2009, pp. 2198-2208.
- [10] Salinas E and Sejnowski T J, "Integrate and fire neuron driven by correlated stochastic input", *Neural Computation*, Vol.14, 2002, pp. 2111-2155.
- [11] Tsotsos J K, Culhane S M, Wai W Y K, Lai Y, Davis N and Nuflo F, "Modeling visuall attention via selective tuning", *Artificial Intelligence*, Vol 78, 1995, pp. 507-545.
- [12] Tyagi A., Yadav A. and Kumar S., "Information processing in brain modeling: challenges and opportunities, *IJCA*, pp. 27-31, 2014.
- [13] Vikram T N, Tscherepanow M and Wrede B, "Impact of real time visual attention on computer vision products and cognitive robotics", In *Proceedings of the European Future Technologies Conference and Exhibition*, 2011, pp. 332-333.
- [14] Walther D, Rutishauser U, Koch C and Perona P, "selective visual attention enables learning and recognition of multiple objects in cluttered scenes", *computer vision and image understanding*, 2005, pp. 41-63
- [15] Burkitt A N, "A Review of the integrate and fire neuron model: homogeneous synaptic input", *biological cybernetics*, 2006, DOI 10.1007/s00422-006-0068-6.
- [16] Giraudo M T and Sacerdote L, "Simulation methods in neuronal modelling", 1998, pp. 77-83.
- [17] Sun Y and Fisher R, "Object-based visual attention for computer vision", *Artificial intelligence*, pp. 77-123.