A Biologically Plausible Neuromorphic System for Object Recognition and Depth Analysis

Zhijun Yang, Alan F. Murray

IMNS, School of Engineering and Electronics
The University of Edinburgh
Edinburgh, EH9 3JL, UK
Emails: {Zhijun.Yang, Alan.Murray}@ee.ed.ac.uk

Abstract.   We present a large-scale Neuromorphic model based on integrate-and-fire (IF) neurons that analyses objects and their depth within a moving visual scene. A feature-based algorithm builds a luminosity receptor field as an artificial retina, in which the IF neurons act both as photoreceptors and processing units. We show that the IF neurons can trace an object's path and depth using an adaptive time-window and Temporally Asymmetric Hebbian (TAH) training.

1   Introduction

Object recognition and the extraction of distance between observer and the object, i.e., the depth information, are closely related in cortical computation. However, the projection of a 3-dimensional scene on to a 2-dimensional focal plane (FP) leads to the loss of depth information (in the third dimension) unless mechanisms are put in place to retain or reconstruct it. Various methods have therefore been introduced to explore the depth recovery ability of creatures under such a 3D-2D projection [1][2][3]. In a dynamic scene, for example, an optic flow field is generated by object motion or by ego-motion of the observer and is projected continuously on to the FP. The optic flow field is thus a re-representation of the spatiotemporal information in the original 3-dimensional space and has many cues from which the original attributes of an object can be reconstructed.

In a sampled system, the flow field on a FP is series of discrete image frames, densely sampled during object/ego motion. Our feature-based object recognition algorithm therefore relies upon a previous frame(s) to predict the re-occurrence of a particular feature in the FP. A causal relationship is thus implicit in the operation of the algorithm. Recent neurobiological results have shown that Temporally Asymmetric Hebbian (TAH) learning has its roots in causality. It is believed that TAH processes drive synaptic plasticity in diverse brain regions to coordinate pre- and postsynaptic

neuronal activities within a critical time window and thus effect information storage and processing[4][5][6].

The integrate-and-fire (IF) neuron is a simple, nonlinear and analog model of the real neuron and is amenable to analogue VLSI (aVLSI) implementation. We therefore use IF neurons to form a spike-processing network for object recognition and depth analysis, based upon the chosen feature-based algorithm[9]. After TAH adaptation of the neurophysiological parameters, IF neurons have a "firing window" adapted to its environmental uncertainty and intrinsic inaccuracy (e.g. circuit imperfection). Adaptation of the width of the firing window is implemented by the TAH rule. By incorporating adaptation this novel IF-based network aims to be a neuromorphic system capable of emulating some functionalities of retina and visual cortex.

2 Methodology

2.1 Optical and neurobiological description

Recently Wörgötter et al. [9] devised a feature-based approach to the recovery of depth information from radial flow fields. The two key steps in the algorithm are: (1) calculation of the object coordinates in the environment; (2) a predictive mechanism to compare the predicted and actual time of arrival of an edge feature for pattern recognition and to improve noise rejection. A temporal tolerance or "window" is associated with neurons, which are arranged radially in the focal plane. At an individual pixel/neuron, a previously calculated edge is accepted if the predicted and actual arrival time of that edge feature are within the neuron's tolerance window. Otherwise the edge coordinates are rejected as spurious. We implement this mechanism as a spike-based scheme and by doing so we can design an artificial vision device through the phenomenological model, which is also amenable to aVLSI implementation.
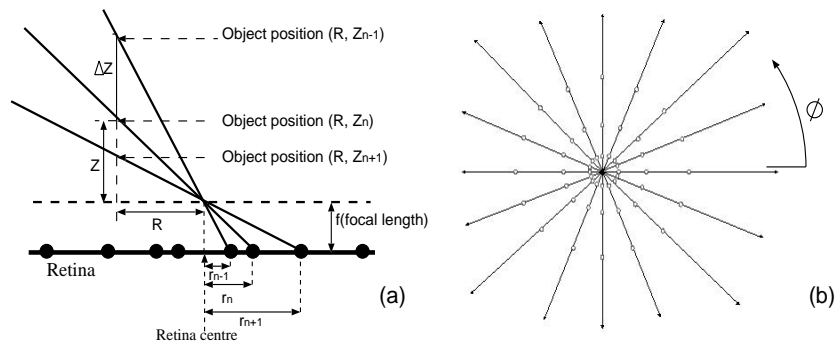

Figure 1: The geometrical projection of a point on to a retina through the light rays (side view), the dark circles are neurons (a), and the schematic layout of a retina, the small circles are neurons (top view) (b).

We consider a dynamic scene in which an observer moves along the optical axis

towards the target objects at a constant speed. The retina has photoreceptor neurons aligned along the radially arranged axes. The neurons on an axis are separated by a hyperbolically increasing distance $d_n = sinh(\frac{n}{\alpha})$, where n is the neuron order and α is a constant, chosen empirically according to be suitable for the pixel resolution. For a suitable α value, higher photoreceptor density can be achieved by interpolating additional neurons between the already arranged ones. The distance from each neuron to the retina centre is a function of its order referred to as $r_n = g(n)$.

The derivation of the depth information for a neuron is conceptually straightforward [9]. In cylindrical polar coordinates we have,

$$\begin{pmatrix} R \\ Z \end{pmatrix}_\phi = \Delta Z \begin{pmatrix} \frac{r_n}{f(k_n-1)} \\ \frac{1}{k_n-1} \end{pmatrix}_\phi \tag{1}$$

where R, Z, ϕ are the cylindrical polar coordinate components determining the actual position of an object in a $3-$D space, $k_n = \frac{r_n}{r_{n-1}}$, f is the focal length, ΔZ is the previous moving distance with projection from neuron $n-1$ to n. It is obvious that, except ΔZ, the right hand side of this polar coordinate equation is fixed for each neuron. Assuming that the ego/object motion speed is constant and known, then ΔZ is available through a step counter to compute the time a light ray needs to travel between the corresponding neurons, the depth information is hence readily recovered.

2.2 The neuronal model

Each neuron in the artificial retina network is a leaky IF unit with two identical excitatory synapses configured to accept pulsed or "spiking" inputs. The synapses connecting the previous neighbouring neuron and the photoreceptor are referred to as the *flow synapse* and the *receptive synapse*, respectively. For analytical simplicity we make 2 assumptions. (i) No delay is associated with input pulse propagation within the axons. (ii) The IF neuron fires only in response to excitatory postsynaptic potentials (EPSPs) induced by a pair of synaptic inputs from the two different synapses. No individual input pulse or consecutive pulses from a same synapse can activate the neuron.

The sub-threshold membrane status when such a pair is presented is:-

$$\begin{cases} \tau_{mem} \frac{dv(t)}{dt} = v_m - v(t) + r_m \sum_i I_i^{syn} \\ \tau_{syn} \frac{dI_i^{syn}(t)}{dt} = -I_i^{syn}(t) + w_i \sum_j \delta(t-t_j) \end{cases} \tag{2}$$

where τ_{mem} and τ_{syn} are the membrane and synaptic time constants respectively, r_m is the membrane resistance, w_i is the weight of a synapse, $\delta(t-t_j)$ is a δ-function dendritic input at time t_j, $I_i^{syn}(t)$ is the current pulse at the post-dendrite terminal, $v(t)$ is the EPSP induced by $I_i^{syn}(t)$, in our case $i = 1, 2$.

The firing threshold of an IF neuron is a combination of an exponentially decaying threshold $V(t) = ae^{-\frac{t}{\tau_{th}}}$ and an intersecting constant threshold $V(t) = V_T$. The use of such a dynamic threshold for pattern formation is not biologically implausible[8]. The threshold equation is then:-

$$V_{th} = \begin{cases} ae^{-\frac{t}{\tau_{th}}} & \text{if } 0 \leq t \leq -\tau_{th} \ln \frac{V_T}{a} \\ V_T & \text{if } -\tau_{th} \ln \frac{V_T}{a} < t \end{cases} \qquad (3)$$

where a is the threshold value at $t = 0$, the time when the previous neighbouring neuron fires, τ_{th} is the exponential threshold decay rate, V_T is the constant threshold value. $-\tau_{th} \ln \frac{V_T}{a}$ is a time instant when the predicted feature arrives at the neuron. If the neuron is depolarised by its *flow synapse* at time t, we would expect it to be further depolarised by its *receptive synapse* in a time range of $[t - \tau_{th} \ln \frac{V_T}{a} - \Delta t, t - \tau_{th} \ln \frac{V_T}{a} + \Delta t]$, where Δt is the response window to be determined by the adaptation mechanism. If the depolarisation from the *receptive synapse* is within the expected response window then the neuron issues a spike and its membrane potential is reset.

2.3 Adaptation mechanism

We aim to make this neuromorphic reformulation of the depth-recovery algorithm adaptive, and thus robust against inevitable inaccuracy in circuits and fabrication processes. We use the TAH learning rule, which responds to spike synchrony, or near-synchrony. The adaptation of each neuron in the artificial retina potentiates or depresses the *receptive synapse* weight by a small amount according to whether the depolarisation from the *receptive synapse* is before or after the predicted time instant. The *flow synapse* weight is held constant for the model simplification. The adaptation scheme is:-

$$f(w) = \begin{cases} A_+ e^{\frac{t-t_{pred}}{\tau_+}} & \text{if } t < t_{pred} \\ -A_- e^{-\frac{t-t_{pred}}{\tau_-}} & \text{if } t > t_{pred} \\ 0 & \text{if } t = t_{pred} \end{cases} \qquad (4)$$

where f(w) is the synaptic weight change, $\{A_+, \tau_+\}$ and $\{A_-, \tau_-\}$ are the initial amplitude and decay constant of potentiation and depression, respectively and t_{pred} is the predicted time instant. In an initial stage, the untrained weight of each neuron is set to a value corresponding to a very narrow window size, suitable for imperfection-free circuitry only. This idealised set-point corresponds to the origin in the TAH learning curve. As the neural and pixel circuits are not, of course, actually ideal, the TAH process will modify the window size.

3 Simulation results

This spike-based algorithm is tested with an artificial environment in which there are three objects (cylinder, cone, sphere) located at different distances ($10m$, $9m$, $7m$) in front of a white background $11m$ away. The observer is 1.6m high and moves towards the scene. In this experiment, object edges are used as the stimulus to the IF neurons and depth maps are defined only at the edges of the objects. The test parameters are shown in table 1. All the neuron parameters except the initial amplitudes of learning curves $A_+ = 0.2$ and $A_- = -0.3$ are chosen with guidance from neurobiology. We

choose A_+ and A_- to be larger than those ones used in an earlier study [7], as our pixel stimuli are much sparser. We use Euler integration as the numerical method and the update step is increased with each image frame.

Scene-Retina Parameter	Value	Neuron Parameter	Value
Sequence	600 frames	Membrane τ	20ms
Resolution	512x512 pixel	Synapse τ	5ms
Step	1cm/frame	Adapt curve τ	20ms
Axes number	400	Rest potential	-70mv
Neuron/Axis	48	Reset potential	-70mv
Retina radius	250 pixel	Constant threshold	-56mv

Table 1: Simulation parameters for object recognition and depth recovery

The depth map is shown in fig.2. The three objects are identified clearly by the edge features flowing along the retina axis. At the outset, (fig.2a), all features that fall on the photoreceptors are new and consequently stimulate the neighbour neurons via their *flow synapses*. Few edges appear on the depth map at this stage.

After this initial settling period, as the actual light rays move from the original photoreceptors to the next ones, the stimuli from the *receptive synapses* begin to depolarise the neurons' membrane.

If the EPSPs from the two synapses to a neuron sum within the firing window to produce an activity that exceeds than the combined threshold, then the neuron is activated to confirm the identification of the associated edge feature. The confirmed feature is then included in the depth map. Meanwhile its depth information can be calculated by the neuron according to equation (1). As further movement occurs in the real scene and the confirmed feature continues to flow along the retinal axis it can be confirmed repeatedly to become an ever more reliable representation of the edge that it represents. In the depth map in fig.2(f), the geometrical shape of objects is clear and depth information can be recovered as described above.

4 Conclusion

We have developed an artificial retina in simulation, which is able to recognise objects and to recover associated depth information using spiking computation and a temporally-adaptive algorithm to optimise performance in the presence of inevitable system and sensor imperfections. The experiments show that, as the system imperfections are fixed by the retinal layout, the *receptive synapse* weights converge to the upper or lower limits of the weight range. This result is consistent with Wörgötter et al.'s layout error analysis [9]. In summary, we do not claim to have developed the optimal depth analysis methodology. Rather, we have shown that neuromorphic and biologically plausible components, together with spike-timing-based adaptation, can compensate for imperfections in a depth-sensitive retina.
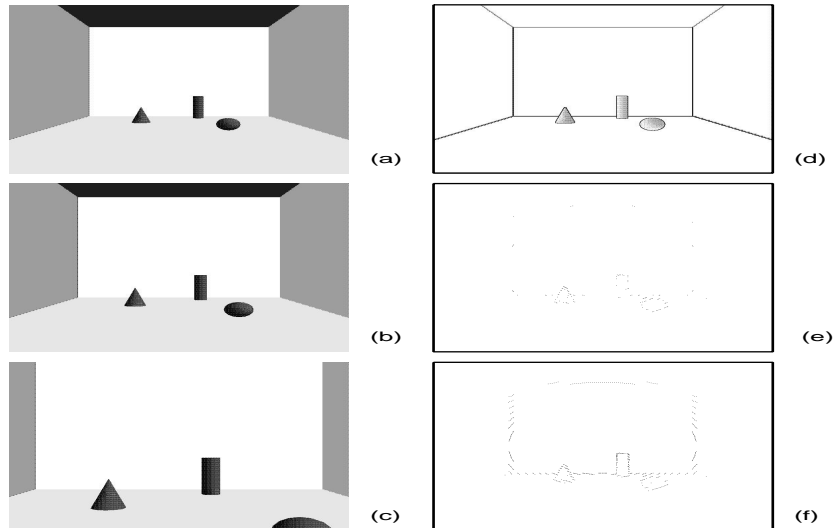
Figure 2: The depth maps of a hallway scene. $(a) \sim (c)$ are the snapshots of image frames at the start, 175, and 600 steps respectively. (d) is the still contour of the first frame without running the spike-based algorithm. (e) and (f) are the depth maps calculated from the start frame till the corresponding frame shown on the left with the algorithm.

References

[1] Qian, N., *Binocular disparity and the perception of depth*, Neuron **18**:359-368, 1997.

[2] Zhou, C. & Tao, H., *Dynamic depth recovery from unsynchronized video streams*, Proc. IEEE Conf. on CVPR, 2003.

[3] Barron, J.L., Fleet, D. & Beauchemin, S., *Performance of optical flow techniques*, Int. J. Comp. Vis. **12**:43-77, 1994.

[4] Markram, H., Lubke, J., Frotscher M. & Sakmann B., *Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs*, Science **275**:213-215, 1997.

[5] Bi, G.Q. & Poo, M.M., *Activity-induced synaptic modifications in hippocampal culture*, Journal of Neuroscience **18**:10464-10472, 1998.

[6] Zhang, L.I., et al., *A critical window for cooperation and competition among developing retinotectal synapses*, Nature **395**:37-44, 1998.

[7] Song, S., Miller, K.D. & Abbott, L.F., *Competitive Hebbian learning through spike-timing-dependent synaptic plasticity*, Nature Neuroscience **3(9)**:919-926, 2000.

[8] Perkel, D.H., *A computer program for simulating a network of interacting neurons. I. Organization and physiological assumptions*, Comput. Biomed. Res. **9**:31-43, 1976.

[9] Wörgötter, F., Cozzi, A. & Gerdes V., *A parallel noise-robust algorithm to recover depth information from radial flow fields*, Neural Computation **11**:381-416, 1999.