Biologically Inspired Edge Detection

Dermot Kerr, Sonya Coleman, Martin McGinnity, QingXiang Wu
Intelligent Systems Research Centre,
University of Ulster Magee,
Derry, U.K., BT48 7JL.
{d.kerr, sa.coleman, tm.mcginnity, q.wu}@ulster.ac.uk

Marine Clogenson
CPE Lyon,
Domaine Scientifique de la Doua,
BP 82077-69616 Villeurbanne, France
marine.clogenson@cpe.fr

***Abstract*—Inspired by the structure and behaviour of the human visual system, we present an approach to edge detection using spiking neural networks and a biologically plausible hexagonal pixel arrangement. Standard digital images are converted into a hexagonal pixel representation and then processed using a spiking neural network with hexagonal shaped receptive fields. The performance is compared with receptive fields implemented on standard rectangular images. Results illustrate that, using hexagonal shaped receptive fields, performance is improved over standard rectangular shaped receptive fields**

Keywords-component; Spiking neural network; Edge detection

I. INTRODUCTION

The human visual system displays very powerful biological processing functionalities that traditional computer vision techniques have not yet fully emulated. Biological research has shown that the brain deals with information processing by using a complicated network of neurons [1]. The process of simulating biological information processing in engineering is termed neuro-engineering [2] and such techniques are typically used for artificial intelligent systems. However, as precise knowledge of the complete neuron circuits in the brain is still not available, it is difficult to implement detailed exact models of biological processing. Thus, most current artificial models are based on specific assumptions and simplified biological processes.

The human visual system is an intrinsically complex hierarchical processing system with various layers and feedback loops. The first stage of visual processing begins in the retina where typically 6 million cone photoreceptors (arranged in a hexagonal lattice) and 90 million rod photoreceptors transform visible light into chemical signals. The signals are then processed by approximately 55 distinct cell types in the retina, arranged in various layers performing different processing functions [3]. The result from this processing is 1.2 million retinal ganglion cell axons conveying the visual scene using action potentials (or spikes) along the optic nerve to the lateral geniculate nucleus and onwards to the various layers of the visual cortex. Various types of receptive fields have been identified [4]; within the retina simple on-off centre-surround receptive fields have been identified that respond to contrast change [4] and similarly these receptive fields have been identified in the lateral geniculate nucleus. Within the visual cortex more complex receptive fields have been identified [4, 5] where each type responds to different stimuli, for example orientated features.

While the cone photoreceptors found in a biological vision system, such as the human retina, are typically arranged in a hexagonal lattice, most existing methods of computer vision are based on conventional rectangular lattices. Previous research [6] has shown that curved structures are not well represented on a rectangular lattice leading us to question why we use them when nature has chosen a hexagonal lattice for human photoreceptors? Using an artificial hexagonal sampling lattice both spatial and spectral advantages may be derived: namely, equidistance of all pixel neighbours and improved spatial isotropy of spectral response. Pixel spatial equidistance facilitates the implementation of circular symmetric kernels that are associated with an increase in accuracy when detecting edges, both straight and curved [6]. Additionally, better spatial sampling efficiency is achieved by the hexagonal structure compared with a rectangular grid of similar pixel separation, leading to improved computational performance. In a hexagonal grid with unit separation of pixel centres, approximately 13% fewer pixels are required to represent the same image resolution as required on a rectangular grid with unit horizontal and vertical separation of pixel centres [7].

In this paper we present an approach to biologically inspired edge detection by using spiking neural networks to more accurately mimic the biological information processing in the brain. A biological plausible hexagonal image arrangement, similar to is the hexagonally arranged photoreceptors found in the retina is used, with hexagonally arranged near-circular receptive fields that emulate those found in the human visual system.

II. CREATING THE HEXAGONAL IMAGE

There is currently no commercially available hardware to capture or display hexagonal images and therefore a resampling technique must be applied to generate hexagonal pixel based images. Many resampling techniques exist for this purpose; here we use the technique proposed in [9] in which Middleton enhances Wuthrich's [10] method of creating a pseudo hexagonal pixel from a cluster of square

978-1-4577-1676-8/11/$26.00 ©2011 IEEE

pixels by representing each pixel by a pixel block in order to create a sub-pixel effect which enables the sub-pixel clustering; this limits the loss of image resolution whilst complying with the main hexagonal properties. Selection of the number of pixels to be clustered for each hexagonal pixel is based on two issues: the arrangement must allow a tessellation with no overlap and no gaps between neighbouring hexagonal pixels; and the cluster must closely resemble a hexagon i.e. six sides of approximately equal length. In [10], two possible choices of hexagonal pixel representations are presented: in one case the hexagonal pixel is comprised of 30 sub-pixels, in the other case it is comprised of 56 sub-pixels; we have chosen to use the 56 sub-pixel approach as illustrated in Fig. 1. To avoid a high loss in image resolution when using this technique each original pixel is separated into a 7×7 block of sub‐pixels having the same intensity as the original pixel. Each hexagonal pixel is then created by clustering 56 of thesesub‐pixels together with its intensity being calculated as the average intensity of the 56 sub‐pixels. The image resizing also enables the display of sub pixels, and therefore the display of hexagonal pixels. When comparing the size of these generated hexagonal hyper‐pixels with the original square pixels, the hexagonal pixels are 12.5% larger than the square pixels. Thus the resampled image has the same resolution as the original image but the number of hexagonal pixels necessary to tessellate an image plane is 12.5% less than it would take using square pixels. With this structure now in place, a cluster of sub pixels in the new image, closely representing the shape of a hexagon, can be created that represents a single hexagonal pixel in the resized image.

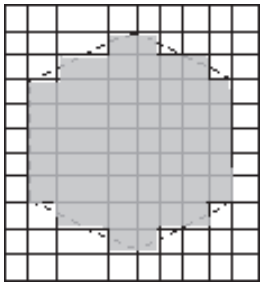


Figure 1. 56 sub-pixel cluster.

III. SPIKING NEURON MODEL

The most widely used spiking neuron model was developed from Hodgkin and Huxley's work [11] based on experimental recordings obtained from experiments on the giant squid axon using a voltage clamp method. However, even though this model is biologically plausible, the complexity in simulating the model is very high due to the number of differential equations. Thus, most computer simulations of neuron models choose to use a simplified neuron model such as the integrate-and-fire model (I&F), leaky I&F model, conductance-based I&F or Izhikevich's model. A full review of the biological behaviour of single neurons can be found in [12] and a comparison of different neuron models can be found in [13]. For implementation purposes the conductance-based I&F model has been selected to model the network neurons in this work. This model offers similar neuron behaviour to the Hodgkin-Huxley whilst providing a reduction in computational complexity. In the conductance-based I&F model the membrane potential $v(t)$ is governed by the following equation:

$$c_m \frac{dv(t)}{dt} = g_l(E_l - v(t)) + \frac{w_{ex} g_{ex}(t)}{A_{ex}}(E_{ex} - v(t)) + \frac{w_{ih} g_{ih}(t)}{A_{ih}}(E_{ih} - v(t)) \quad (1)$$

where c_m is the membrane capacitance, E_l is the membrane reversal potential, g_l is the conductance of the membrane, E_{ex} and E_{ih} are the reversal potential of the excitatory and inhibitory synapses respectively, w_{ex} and w_{ih} are weights for excitatory and inhibitory synapses respectively, and A_{ex} and A_{ih} are the membrane surface area connected to the excitatory and inhibitory synapses respectively. If the membrane potential $v(t)$ exceeds the threshold voltage v_{th} an action potential is generated and then $v(t)$ is reset to v_{reset} for a time τ_{ref} which is called the refractory duration. For simplicity τ_{ref} is set to 0 in this paper. The variables $g_{ex}(t)$ and $g_{ih}(t)$ represent the conductance's of excitatory and inhibitory synapses respectively, which vary with time. The output spike train is then represented by a series of 1s or 0s representing whether or not a neuron fires at time *t*, i.e. $[S_{out}(t_1), S_{out}(t_2), \ldots, S_{out}(t_M)]$.

IV. SPIKING NETWORK STRUCTURE & IMPLEMENTATION

In a biological system a receptive field is where a spiking neuron integrates the spikes from a group of afferent neurons as illustrated in Fig. 2 where neuron N has a receptive field with a 7-neuron hexagonal array. Each neuron in the receptive field connects to neuron N through both excitatory and inhibitory synapses.

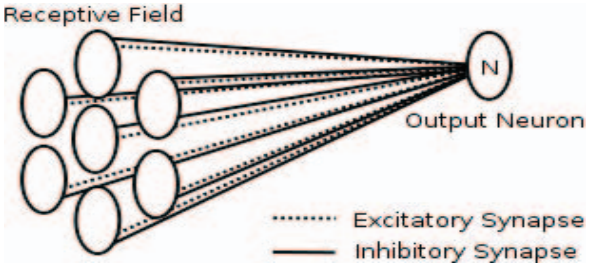


Figure 2. Receptive field of a spiking neuron.

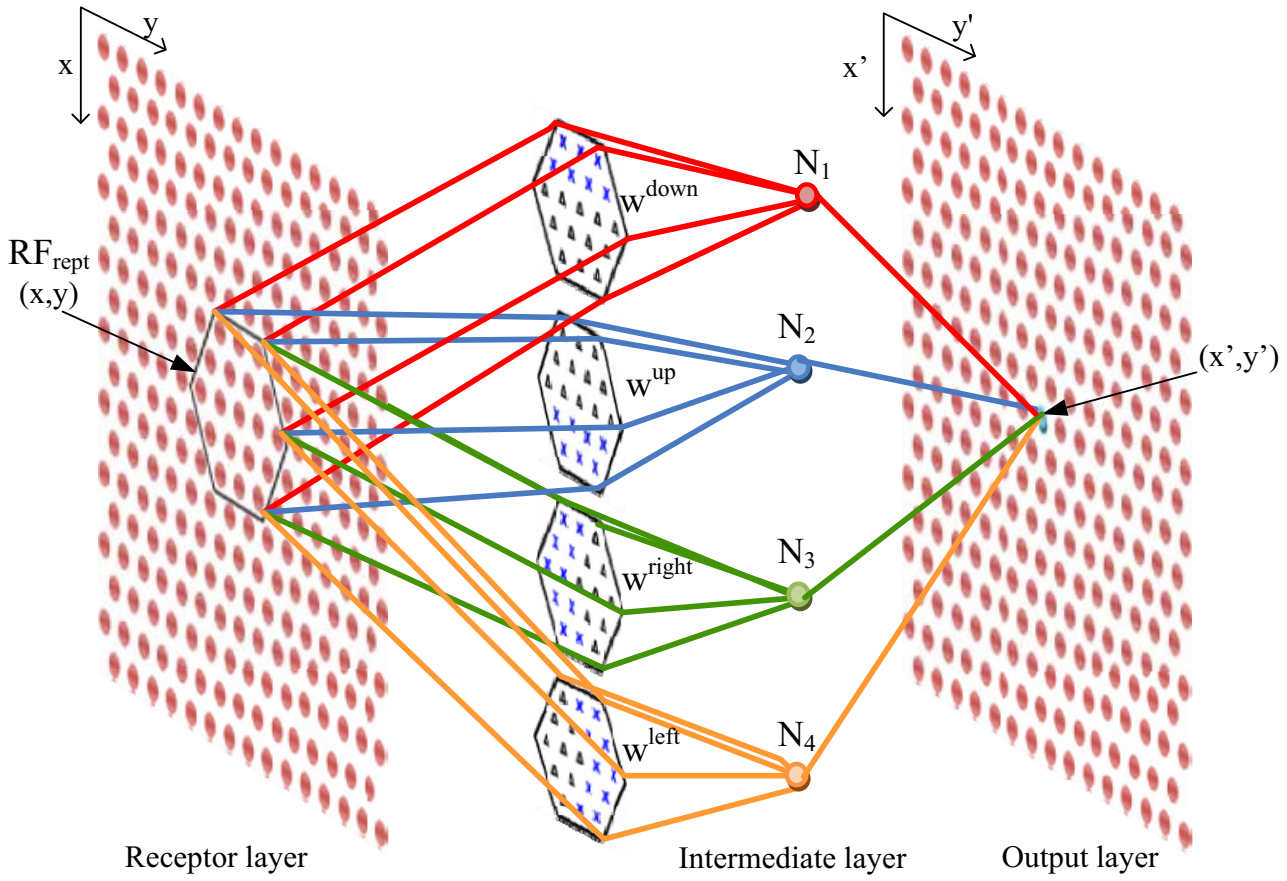


Figure 3. Spiking Neural Network Structure.

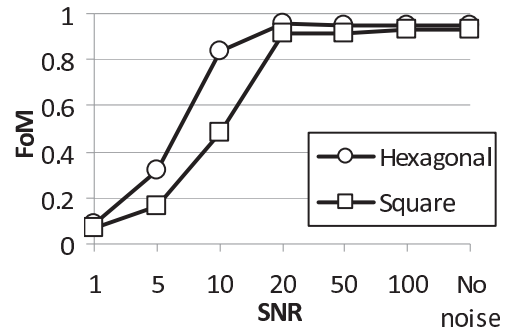
Within the network structure proposed we use four types of receptive fields corresponding to different edge directions using the spiking neuron model described in Section 3. We define our spiking neural network structure as illustrated in Fig. 3.

Suppose that the first layer in Fig. 3 represents photoreceptors. Each pixel in the hexagonal image corresponds to a photoreceptor. The intermediate layer is composed of four types of neurons corresponding to four different receptive fields respectively. 'X' in the synapse connections represents an excitatory synapse. 'Δ' represents an inhibitory synapse. Each neuron in the output layer integrates four corresponding outputs from the intermediate neurons. The firing rate map of the output layer forms an edge graphic corresponding to the input image. There are four parallel arrays of neurons in the intermediate layer each of the same dimension as the receptor layer with only one neuron in each array illustrated in Fig. 3 for simplicity. Each of these intermediate neurons performs the processing for different edge directions and is connected to the receptor layer by differing weight matrices. These weight matrices can be of varying sizes to represent the width of the receptive field under consideration although at this stage we are concerned only with a fixed size. The weights are calculated using the function provided in [8] and, for example, the 19-point hexagonal weight matrices for top and bottom edges are defined as:

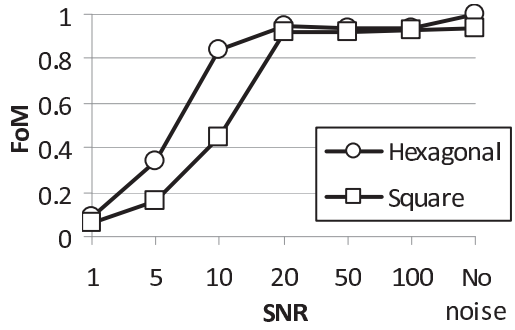
$$\begin{bmatrix} & 0.3548 & & 0.3679 & & 0.3548 & \\ 0.9216 & & 0.9910 & & 0.9910 & & 0.9216 \\ 0 & & 0 & & 0 & & 0 & & 0 \\ & 0 & & 0 & & 0 & & 0 & \\ & & 0 & & 0 & & 0 & & \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} & & 0 & & 0 & & 0 & & \\ & 0 & & 0 & & 0 & & 0 & \\ 0 & & 0 & & 0 & & 0 & & 0 \\ 0.9216 & & 0.9910 & & 0.9910 & & 0.9216 & & \\ & 0.3548 & & 0.3679 & & 0.3548 & & & \end{bmatrix} \quad (3)$$

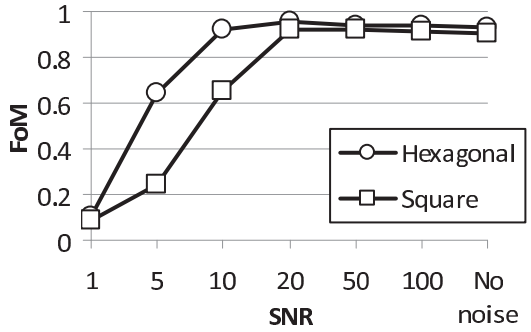
The network model was implemented in Matlab using the same network parameters as found in [8] that are consistent with biological neurons [3]. No training is used in this network and synaptic strengths are adjusted heuristically to ensure that the neuron does not fire in response to a uniform image within its receptive field and the image gray scale values are normalised to a real number in the range $[0,1]$.



(a) Vertical edge



(b) Horizontal edge



(c) Diagonal edge

Figure 4. Figure of Merit results for 19-point hexagonal and 5×5 square receptive fields

V. Experimental Results & Evaluation

To evaluate the edge detection performance of the SNN based approach we have chosen to use the Figure of Merit (FoM) proposed by Pratt [14]. This measure balances three types of error associated with the determination of an edge: missing valid edge points; failure to localise edge points; classification of noise fluctuations as edge points. The FoM is defined as:

$$R = \frac{1}{\max(I_A, I_I)} \sum_{i=1}^{I_A} \frac{1}{1+\alpha d^2} \quad (4)$$

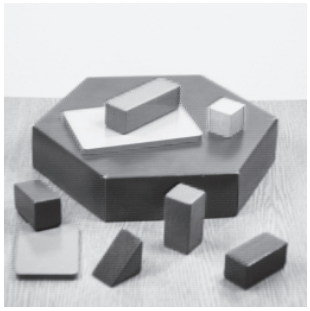
Here I_A is the actual number of edge pixels detected, I_I is the ideal number of edge pixels, d is the separation distance of a detected edge point normal to a line of ideal edge points, and α is a scaling factor, most commonly chosen to be 1/9, although this value may be adjusted to penalise edges that are localised but offset from the true edge position. In Fig. 4 we present evaluation of the hexagonal SNN based edge detector compared with the square SNN based edge detector presented in [8] using various edge orientations. The receptive field of the square edge detector is size 5×5 and correspondingly we use a 19-point hexagonal operator. The Figure of Merit [14] is compared over a range of signal to noise levels (SNR). Fig. 4 illustrates that the hexagonal receptive fields shows improved performance over the square based receptive fields for all edge types. In particular we note that in areas of high noise the hexagonal receptive field performs significantly better that the equivalent square receptive field.

The simulation is run where spikes are computed over a time interval of 100ms. We have computed the time to run this simulation as 3.4687s on the hexagonal arrangement and 3.9219s on the square arrangement, illustrating a slight improvement in computation time with the hexagonal arrangement.

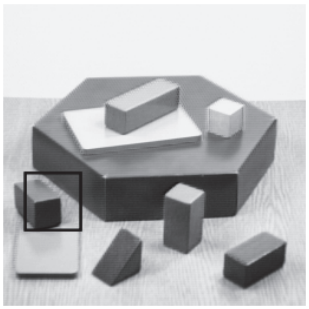
In Fig. 5 we present the 'blox' image in Figure 6 we present the Lena image, which are the inputs to the network, and illustrate example outputs for both the hexagonal and square receptive fields. For illustration purposes, Fig. 5(a) and Fig. 6(a) are hexagonal pixel-based images, however it should be noted that the hexagonal receptive fields are applied only to hexagonal pixel-based images and the square receptive fields are applied only to standard original square pixel based images. In these images, as the edge brightness increases the firing rate of the neuron becomes stronger, thus the firing rate may be set as a threshold to determine the presence or absence of an edge. It can be seen from Fig. 5(c), Fig. 5(e), Fig. 6(c) and Fig. 6(e) that the output from the hexagonal receptive field is much clearer and less noisy than the corresponding output from the square. This is further highlighted examining the corresponding zoomed regions in Fig. 5(d), Fig. 5(f), Fig. 6(d) and Fig. 6(f).

VI. Discussion & Future Work

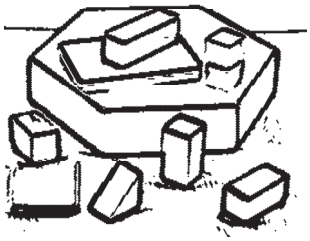
The spiking neural network presented in this paper is constructed by a hierarchical structure that is composed of spiking neurons with various receptive fields. The input image has a hexagonal pixel arrangement and the receptive fields used are arranged in a hexagonal structure. The spiking neuron models provide powerful functionality for integration of inputs and generation of spikes. Synapses are able to perform different complicated computations. This paper demonstrates how a spiking neural network can detect edges in an image using a hexagonal structure and illustrates performance and computational improvements over the standard square based approaches.



(a) Blox input image

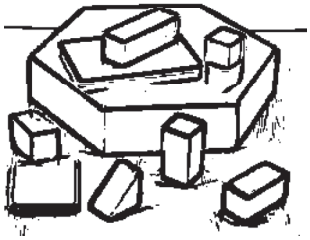


(b) Zoomed region from (a) highlighted



(c) Hex-SNN output

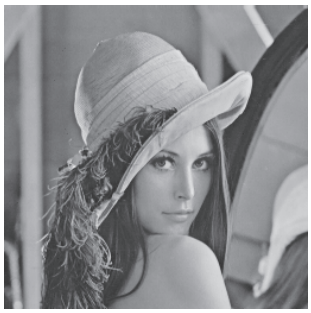
(d) Hex-SNN output zoomed to region highlighted in (b)



(e) Square-SNN output

(f) Square-SNN output zoomed to region highlighted in (b)

Figure 5. Network inputs and outputs using blox image



(a) Lena input image

(b) Zoomed region from (a) highlighted

(c) Hex-SNN output

(d) Hex-SNN output zoomed to region highlighted in (b)

(e) Square-SNN output

(f) Square-SNN output zoomed to region highlighted in (b

Figure 6. Network inputs and outputs using lena image

ACKNOWLEDGMENT

This work was supported by the Centre of Excellence in Intelligent Systems project, funded by InvestNI and the Integrated Development Fund

REFERENCES

[1] Izhikevich, E.M., Edelman, G.M., "A Large-Scale Model of Mammalian Thalamocortical Systems", Proc. National Academic Science, USA, vol. 105, no. 9, pp.3593-3598, 2008.

[2] O'Connor, D.H., Huber, D., Svoboda, K., "Reverse engineering the mouse brain", Nature 461, 923-929, March, 2008.

[3] Masland, R.H., "The fundamental plan of the retina", Nature Neuroscience, vol. 4, pp. 877-886, 2001.

[4] Hosoya, T., Baccus, S.A., Meister, M., "Dynamic predictive coding by the retina", Nature, vol. 436, pp. 71-77, 2005.

[5] Kandel, E.R., Shwartz, J.H., "Principles of neural science", Edward Arnold (Publishers) Ltd., 1981.

[6] Allen J.D., "Filter Banks for Images on Hexagonal Grid," Signal Solutions, 2003.

[7] Vitulli R., "Aliasing Effects Mitigation by Optimized Sampling Grids and Impact on Image Acquisition Chains," Geoscience and Remote Sensing Symposium, pp. 979-981, 2002.

[8] Wu, Q., McGinnity, M., Maguire, L., Belatreche, A., Glackin, B., "Edge Detection Based on Spiking Neural Network Model" Proc Int Conf on Intelligent Computing, LNAI 4682, pp. 26–34, Springer-Verlag Berlin Heidelberg 2007

[9] Middleton L., Sivaswamy, J., "Edge Detection in a Hexagonal-Image Processing Framework," Image and Vision Computing 19, pp. 1071-1081, June 2001.

[10] Wuthrich C.A., Stucki P., "An Algorithm Comparison between Square and Hexagonal Based Grids," CVGIP: Graphical Models and Image Processing 53, pp. 324-339, 1991.

[11] Hodgkin, A., Huxley, A., "A quantitative description of membrane current and its application to conduction and excitation in nerve", Journal of Physiology, London, vol. 117, pp. 500-544, 1952.

[12] Gerstner, W., Kistler, w., "Spiking Neuron Models: Single Neurons, Populations, Plasticity", Cambridge University Press, 2002.

[13] Izhikevich, E.M., "Which model to use for cortical spiking neurons?", IEEE Trans. on Neural Networks, vol. 15, no. 5, 2004.

[14] Abdou, I.E., Pratt, W.K., "Quantitative design and evaluation of enhancement/ thresholding edge detectors" Proceedings of the IEEE, Vol. 67, No. 5, pp. 753-763, 1979.