

A Novel Approach to Robot Vision using a Hexagonal Grid and Spiking Neural Networks

D. Kerr, S.A. Coleman, T.M. McGinnity, Q. Wu
 Intelligent Systems Research Centre
 University of Ulster, Magee
 Londonderry, Northern Ireland, U.K.
 {d.kerr, sa.coleman, tm.mcginny, q.wu}@ulster.ac.uk

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

Abstract— Many robots use range data to obtain an almost 3-dimensional description of their environment. Feature driven segmentation of range images has been primarily used for 3D object recognition, and hence the accuracy of the detected features is a prominent issue. Inspired by the structure and behaviour of the human visual system, we present an approach to feature extraction in range data 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; this approach is a step towards developing a robotic eye that closely mimics the human eye. The performance is compared with receptive fields implemented on standard rectangular images. Results illustrate that, using hexagonally shaped receptive fields, performance is improved over standard rectangular shaped receptive fields.

Keywords-component; range image; hexagonal imaging; spiking neural network

I. INTRODUCTION

In recent years many robotics and computer vision applications have been developed using range image data instead of, or in conjunction with, intensity image data [1]. This is largely because range imagery can be used to obtain reliable descriptions of 3-D scenes; a range image contains distance measurements from a selected reference point or plane to surface points of objects within a scene [2], allowing more information about the scene to be recovered [3]. However, a range image contains information about only the visible surfaces of the objects, and not their hidden surfaces, and hence is often referred to as $2\frac{1}{2}$ -D information [2]. Range images are acquired with range sensors and in an ideal situation, like intensity images, range data are uniformly distributed in the x - and y - directions; however this is seldom the case. A number of range image sensors are available [2], [4], and not all can sample the surface at equidistant x - and y - intervals; often the coordinates of the data points are dependent on the measured range of the point [5], as, for example, in the case of the commonly used ABW, K2T and Perceptron sensors [6], [7] and hence the data are irregularly distributed. Typically range data are interpolated onto a regular grid prior to any further processing. There is no requirement that this regular grid is a rectangular grid and in this paper we propose

to use a hexagonal grid that closely mimics the structure of images captured by the human visual system (HVS) rather than the conventional regular square grid. Curved structures are not well represented on a rectangular lattice and this leads 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 [8].

Cone photoreceptors found in a biological vision system, such as the human retina, are typically arranged in a hexagonal lattice, (as shown in Figure 1) and research has also shown that curved structures are more accurately represented by hexagonal pixels than by rectangular pixels [9]. Recent work that uses the hexagonal structure includes biologically inspired fovea modelling with neural networks [10], and the development of silicon retinas for robot vision [11], [12].

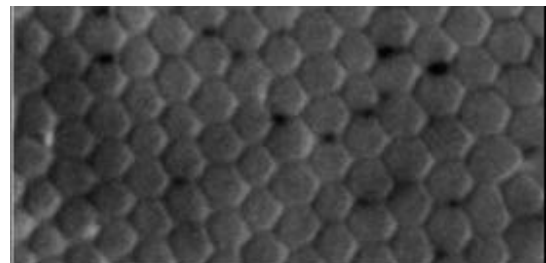


Figure 1. Cross section of human retina showing the hexagonal structure of the photoreceptor cones densely packed within the fovea [13].

Taking additional inspiration from the HVS, research has tried to overcome the failings and computational overhead associated with traditional real-time image processing techniques, typically via the use of neural networks [14]. Spiking neural networks (SNNs) are a class of neural networks that mimic more accurately the biological information processing in the visual cortex, increasing computational power and speed when compared with traditional neural networks and therefore enabling real-time processing [15] which is essential for robotics applications. SNNs use simple neuronal models and communicate using spikes in a manner similar to action

potentials found in biological neurons. There has been some research investigating the application of SNNs to visual processing; a spiking neural network model that performs segmentation and edge detection is proposed in [16]; in [17] a spiking neural network is proposed to detect contours in images through the synchronisation of integrate and fire neurons using simple synthetic images; in [18] a spiking neural network is proposed for real-time edge detection. Additionally, spiking neural networks have been previously used as controllers in evolutionary robotics to perform vision based obstacle avoidance [19], [20], and for laser-based retinal model robot vision [21]. In [22], [23] a robot's sensory information is converted into spikes and a spiking neural network is used to process the information and control the robot. A biologically inspired flying robot is developed in [24] that uses a spiking neural network to convert visual information into motor commands and in [25] a spiking neural network is used to control a mobile robot using sonar sensors. However, none of these algorithms have been developed for range data or can readily be applied to hexagonal image structures.

In this paper we present an approach to biologically inspired feature detection by using spiking neural networks in combination with a hexagonal pixel structure to develop a robotic vision system that reflects a stronger correlation with the human visual system than current systems and has improved performance over the standard use of rectangular pixel images.

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 range images. Many resampling techniques exist for this purpose; here we use is the approach proposed in [26] in which Middleton enhances Wuthrich's [27] method of creating a pseudo hexagonal pixel. In this approach each pixel is represented 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. In [27], 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. The image resizing also enables the display of sub pixels, and therefore the display of hexagonal 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.

III. SPIKING NEURON MODEL

A widely used spiking neuron model is that of Hodgkin and Huxley [28] based on experimental recordings obtained from experiments on the giant squid axon using a voltage clamp method. However, even though this model is biologically

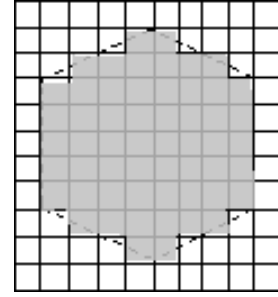


Figure 1. 56 sub-pixel cluster.

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 [29] and a comparison of different neuron models can be found in [30]. 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 areas 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. Then $v(t)$ is then 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. $[S_{out}(t_1), S_{out}(t_2), \dots, S_{out}(t_M)]$.

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.

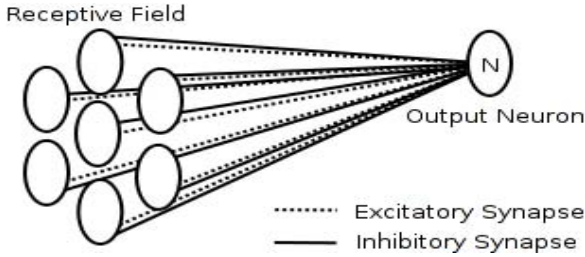


Figure 2. Receptive field of a spiking neuron.

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 III. 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' 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 with the same dimension as the receptor layer. (Only one neuron in each array is illustrated in Fig. 3 for simplicity).

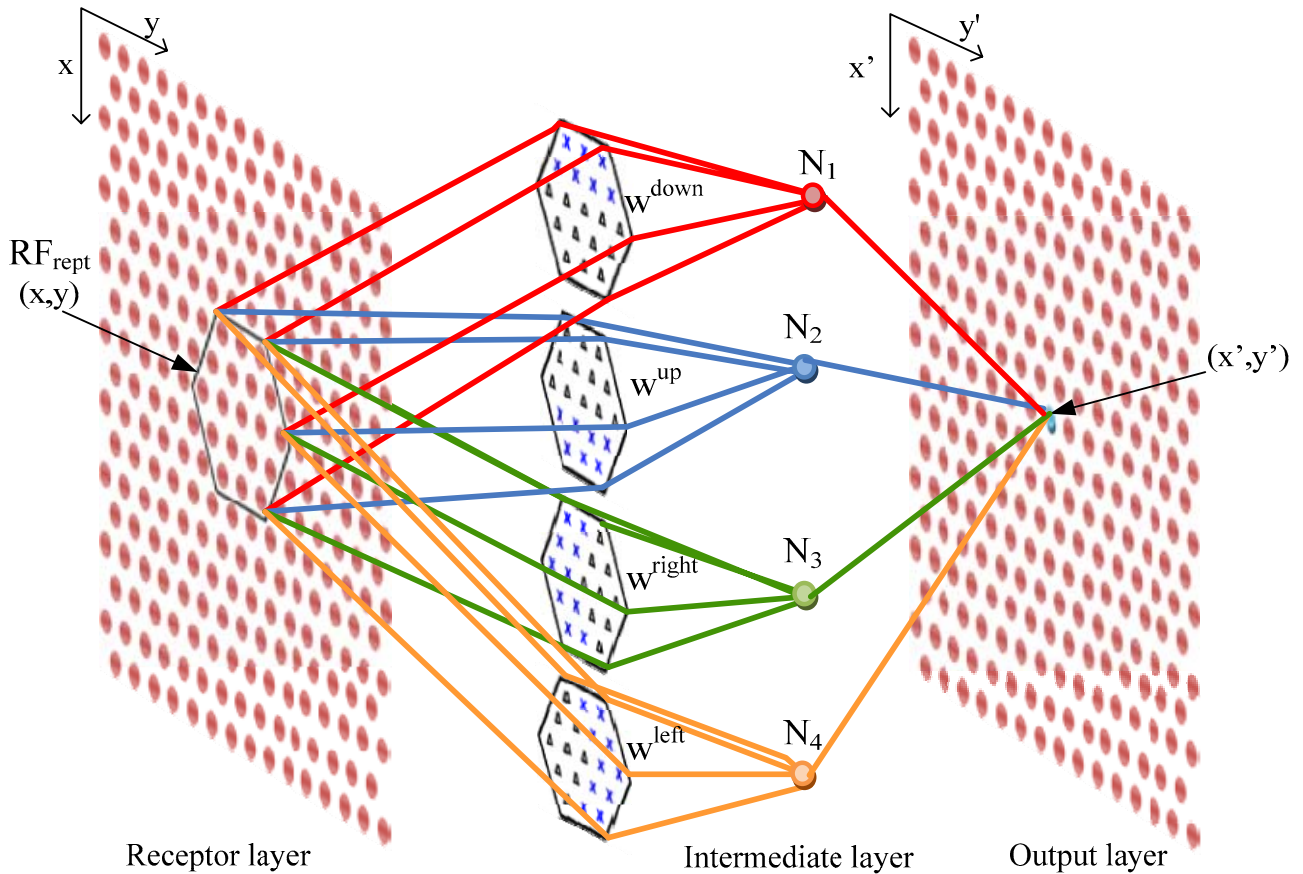


Figure 3. Spiking Neural Network Structure.

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, and in fact we present results for a range of receptive field widths. For a receptive field the weights are calculated using the function provided in [18] and, for example, the 19-point hexagonal weight matrices for top and bottom edges are defined as:

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

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

The network model was implemented in Matlab using the networks parameters found in [18] that are consistent with biological neurons [31]. Synaptic strengths can be adjusted to ensure that the neuron does not fire in response to a uniform image within its receptive field. As previously noted the receptive fields illustrated in the intermediate layer in Fig. 3 can be of any size and in particular we will use 7, 19, and 37-point hexagonal receptive fields.

V. FEATURE EXTRACTION

Here we demonstrate how feature extraction differs when we apply receptive fields to range images, compared to applying such operators to intensity images. Typically with intensity images, after applying gradient operators, thresholding is applied by simply selecting an appropriate threshold value, T , either empirically or scientifically, and all values that lie above T are considered as feature points. However, whilst performing feature extraction on range images we need to consider that the features are represented by depth profiles and features may take the form of significant depth profile changes or depth discontinuities.

The range images used in the experiments are first resampled onto a hexagonal lattice using the method outlined in Section II. The SNN is then constructed in such a manner that the hexagonal structure is maintained through each processing layer. This ensures that the receptive field's

synaptic connections have equidistance of connected neighbours and improved spatial isotropy. Although range images provide the exact spatial (x, y) co-ordinates relating to the depth measurement in space, in this work we use the spatial location closest to the nearest hexagonal pixel using nearest neighbour interpolation. Range image depth values are normalised such that the values are in the range $[0...1]$ and positions in the image with no valid depth measurement are considered to be equal to 0 in order to account for depth discontinuities.

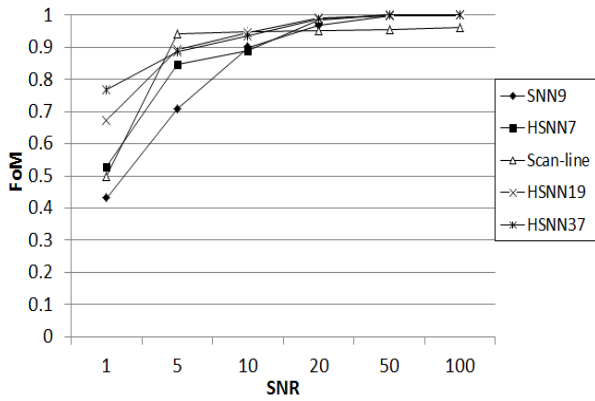
The SNN simulates visual processing though the use of the various receptive fields and as described in Section IV we consider four receptive fields. During the simulation each receptive field is processed simultaneously in time and the output neuron potential is determined by the summation of the combined response from each receptive field. If the output neurons potential reaches the firing threshold during the simulation time, the position is determined as an edge position.

VI. EXPERIMENTAL RESULTS & EVALUATION

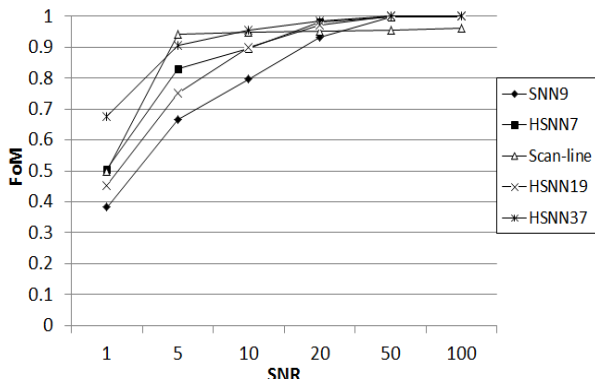
To evaluate the edge detection performance of the SNN based approach we have chosen to use the Figure of Merit (FoM) technique [32]. This technique balances three types of error associated with the determination of an edge: missing valid edge points; failure to localise edge points; classification of 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. We present comparative evaluation for the hexagonal SNN based edge detector the square SNN based edge detector presented in [18] and the well-known scan line approach [6] used for edge detection in range images. The square receptive field is size 3×3 (denoted as SNN9) and we demonstrate results for the hexagonal receptive field at multiple scales: 7-point (equivalent to square 3×3), 19-point and 37 point receptive fields (denoted as HSNN7, HSNN19 and HSNN37). The Figure of Merit [32] is computed over a range of signal to noise levels using both convex and concave roof edge images; two edge types commonly found in range image data. Fig. 4 illustrates that the hexagonal receptive fields shows improved performance over the square based receptive fields for both edge types,



(a) Convex roof edge



(b) Concave roof edge

Figure 4. Figure of Merit results of different edge types over a range of signal-to-noise ratios.

particularly in areas of high noise. Although the scan-line approach performs well in very high noise, it never obtains a FoM value of 1, and therefore it never exactly locates an edge, even when no noise is present.

In Fig. 5(a) we present a range image from the Technical Arts scanner. Fig. 5(b) illustrates the feature map generated using the scan-line approach and similarly Fig. 5(c) – (f) illustrates the outputs from the various SNN based approaches. For illustration purposes, Fig. 5(a) is the original capture image using a standard rectangular grid however it should be noted that this image is resampled on to hexagonal grids for use with the hexagonal receptive fields. 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(d), Fig. 5(e) and Fig. 5(f) that the outputs from the hexagonal receptive field are comparable to the corresponding output from the square receptive field [18] and the scan-line approach [6]. Additional images are presented in Fig. 6.

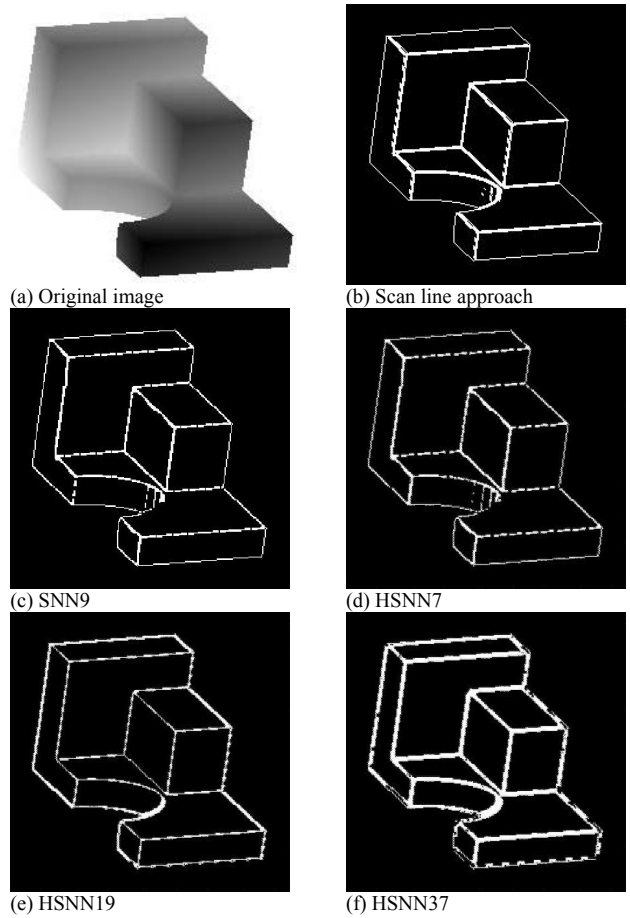


Figure 5. (a) Original image; (b) Feature map generated using the scan-line approach [6]; (c)-(f) Example network outputs using (a) as the input.

Such results are promising as, in addition to the improved successful edge detection performance, there is a potential computational improvement when using a hexagonal grid. The hexagonal grid contains approximately 13% less pixels when compared to a rectangular grid in order to obtain an image of the same spatial resolution; similarly the smallest receptive field contains only 7 values rather than 9 in the standard rectangular receptive field. This computational improvement is illustrated in Table 1 where the time to run the simulation for 100ms is compared illustrating an improvement in computation time with the hexagonal arrangement. Computation time improvements are realised by the decreased hexagonal pixel density in both image and hexagonal operators.

TABLE I. ALGORITHM RUN TIMES (SECONDS).

Receptive field size	Processing time
SNN	3.92
HSNN 7-Point	3.16
HSSN 19-Point	3.47
HSSN 37-Point	3.78

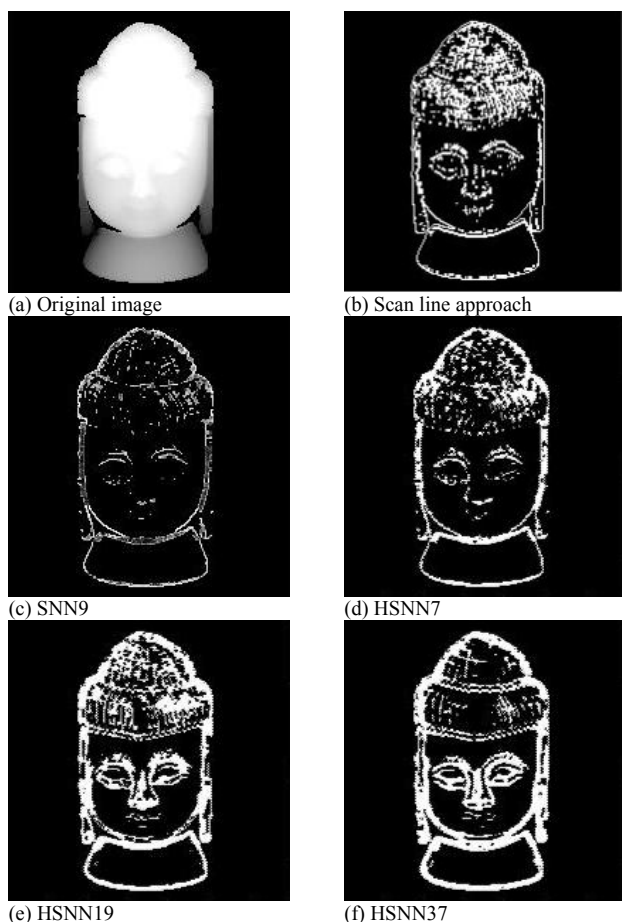


Figure 6. (a) Original image; (b) Feature map generated using the scan-line approach [6]; (c)-(f) Example network outputs using (a) as the input.

VII. DISCUSSION & FUTURE WORK

We present a biologically motivated approach to robotic vision using range image data resampled onto a hexagonal grid and spiking neural networks which compliments recent work in robotic silicon eyes using hexagonal structures[11], [12]. Range images are useful for many robot tasks such as creating models of physical objects or providing information that complements standard intensity images for various robot tasks. Often range data are slightly irregularly distributed and need to be resampled to a regular grid prior to processing and therefore we propose that improved feature extraction results can be obtained by resampling to a hexagonal grid. The input image has a hexagonal pixel arrangement and the receptive fields used are arranged in a hexagonal structure, representing the human fovea. The spiking neural network presented is constructed by a hierarchical structure that is composed of spiking neurons with scalable receptive fields as found in the visual cortex. 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.

ACKNOWLEDGMENT

This research is funded by the Centre of Excellence in Intelligent Systems project, funded by InvestNI and the Integrated Development Fund.

REFERENCES

- [1] P. Dias, et al., "Combining Intensity and Range Images for 3D Modelling", *Proceedings of the IEEE International Conference on Image Processing (ICIP2003)*, 2003.
- [2] P.J. Besl., "Active, optical range imaging sensors." *Machine Vision and Apps*, Vol.1, pp127-152, 1988.
- [3] O. Bellon, and L. Silva, "New Improvements on Range Image Segmentation by Edge Detection Techniques" *Proceedings of the workshop on Artificial Intelligence and Computer Vision*, Nov. 2000.
- [4] R.J. Jarvis, "Range Sensing for Computer Vision", *Three-Dimensional Object Recognition Systems*, Elsevier Science, Amsterdam, pp. 17-56, 1993.
- [5] M. De Bakker, "The PSD chip, high speed acquisition of range images" *PhD Thesis*, Delft University of Technology, 2000
- [6] X. Jiang, and H. Bunke, "Edge Detection in Range Images Based on Scan Line Approximation" *Computer Vision and Image Understanding*, Vol.73, No.2, pp. 183-199, 1999.
- [7] X.J. Jiang, and H. Bunke, "Fast Segmentation of Range Images into Planar Regions by Scan Line Grouping", *Machine Vision and Applications*, 7(2), pp.115-11, 1994.
- [8] He X., Jia W., "Hexagonal Structure for Intelligent Vision," *Information and Communication Technologies, ICICT*, pp. 52- 64, 2005.
- [9] J.D. Allen, "Filter Banks for Images on Hexagonal Grid," *Signal Solutions*, 2003.
- [10] C.H. Huang, C.T. Lin, "Bio-Inspired Computer Fovea Model Based on Hexagonal-Type Cellular Neural Network" *IEEE Trans Circuits and Systems*, 54(1), pp35-47, Jan 2007.
- [11] K. Shimonomura, et al., "Neuromorphic binocular vision system for real-time disparity estimation" *IEEE Int Conf on Robotics and Automation*, pp.4867-4872, 2007.
- [12] R. Takami, et al., "An Image Pre-processing system Employing Neuromorphic 100 x 100 Pixel Silicon Retina" *IEEE Int Symp Circuits & Systems*, Vol. 3, pp2771-2774, 2005.
- [13] C.A. Curcio, et al., "Human Photoreceptor Topography", *Journal of Comparative Neurology*, Vol. 292, pp. 497-523, 1990.
- [14] M. Egmont-Petersen, D. De Ridder. and H. Handels. "Image processing with neural networks-a review". *Pattern Recognition*, 35(10), 2279-2301, 2003.
- [15] D.R. Kunkle, and C. Merrigan. "Pulsed neural networks and their application." *Computer Science Dept., College of Computing and Information Sciences*, Rochester Institute of Technology, 2002.
- [16] B. Meftah, O. Lezoray, and A. Benyettou, "Segmentation and Edge Detection based on Spiking Neural Network Model". *Neural Processing Letters*. 32(2), 131-146, 2010.
- [17] E. Hugues, F. Guilleux, and O. Rochel, "Contour Detection by synchronization of Integrate and Fire Neurons". *Lecture Notes in Computer Science*. 2525, 60-69, 2002.
- [18] Q. Wu, T.M. McGinnity, L.P. Maguire, A. Belatreche, B. Glackin, "Edge Detection Based on Spiking Neural Network Model" *Proc Int Conf on Intelligent Computing*, LNAI 4682, pp. 26-34, Springer-Verlag Berlin Heidelberg, 2007.
- [19] D. Floreano and C. Mattiussi. "Evolution of spiking neural controllers for autonomous vision-based robots." *Evolutionary Robotics IV*, pp 38-61. Springer-Verlag, Berlin, 2001.

- [20] D. Roggen, S. Hofmann, Y. Thoma, D. Floreano. "Hardware spiking neural network with run-time reconfigurable connectivity in an autonomous robot." *Proceedings. NASA/DoD Conference on Evolvable Hardware*, 2003.
- [21] H. Masuta, N. Kubota, "The perception for partner robot using spiking neural network in dynamic environment." *SICE Annual Conference*, 2008.
- [22] D. Gamez, "SpikeStream: A Fast and Flexible Simulator of Spiking Neural Networks." *Proceedings of ICANN*. V.4668, pp. 370-379. Lecture Notes in Computer Science, Springer Verlag, 2007.
- [23] E. Lazdins, A. K. Fidjeland, D. Gamez, "iSpike: A Spiking Neural Interface for the iCub Robot". , *Proceedings of the International workshop on bio-inspired robots*. 2011.
- [24] D. Floreano, J.C. Zufferey, and J.D. Nicoud, "From Wheels to Wings with Evolutionary Spiking Neurons." *Artificial Life*, 11(1-2) pp. 121-138, 2005.
- [25] H. Hagaras, A. Pounds-Cornish, M. Colley, V.Callaghan, and G. Clarke, "Evolving spiking neural network controllers for autonomous robots." *Proceedings of IEEE International Conference on Robotics and Automation*, (ICRA) 2004.
- [26] L. Middleton J. Sivaswamy "Edge Detection in a Hexagonal-Image Processing Framework," *Image and Vision Computing 19*, pp. 1071-1081, June 2001.
- [27] C.A. Wuthrich and P. Stucki P., "An Algorithm Comparison between Square and Hexagonal Based Grids," *CVGIP: Graphical Models and Image Processing 53*, pp. 324-339, 1991.
- [28] A. Hodgkin, and A. Huxley, "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.
- [29] W. Gerstner, and W. Kistler, "Spiking Neuron Models: Single Neurons, Populations, Plasticity", *Cambridge University Press*, 2002.
- [30] E.M. Izhikevich, "Which model to use for cortical spiking neurons?", *IEEE Trans. on Neural Networks*, vol. 15, no. 5, 2004.
- [31] R.H. Masland, "The fundamental plan of the retina", *Nature Neuroscience*, vol. 4, pp. 877-886, 2001.
- [32] I.E. Abdou, and W.K. Pratt, "Quantitative design and evaluation of enhancement/ thresholding edge detectors" *Proceedings of the IEEE*, Vol. 67, No. 5, pp. 753-763, 1979.