

Biologically Motivated Feature Extraction

Sonya Coleman, Bryan Scotney, and Bryan Gardiner

¹ School of Computing and Intelligent Systems, University of Ulster, Magee, UK

² School of Computing and Information Engineering, University of Ulster, Coleraine, UK

{sa.coleman,bw.scotney,b.gardiner}@ulster.ac.uk

Abstract. We present a biologically motivated approach to fast feature extraction on hexagonal pixel based images using the concept of eye tremor in combination with the use of the spiral architecture and convolution of non-overlapping Laplacian masks. We generate seven feature maps “a-trous” that can be combined into a single complete feature map, and we demonstrate that this approach is significantly faster than the use of conventional spiral convolution or the use of a neighbourhood address look-up table on hexagonal images.

Keywords: hexagonal images, feature extraction, spiral architecture, eye tremor.

1 Introduction

In order to obtain real-time solutions to problems that require efficient large-scale computation, researchers often seek inspiration from biological systems; for real-time image processing we consider the characteristics of the human visual system. In order for humans to process visual input, the eye captures information that is directed to the retina located on the inner surface of the eye. A small region within the retina, known as the fovea and consisting of a high density of cones, is responsible for sharp vision capture and is comprised of cones that are shaped and placed in a hexagonal arrangement [4, 6, 8]. Additional important characteristics of the human eye are that, within the central fovea, receptive fields of ganglion cells of the same type do not overlap [5], and that the eye can be subjected to three types of eye movement: tremor, drift, and micro-saccades. Furthermore, the human vision system does not process single static images, but instead a series of temporal images that are slightly off-set due to involuntary eye movements. Therefore, the traditional approaches to feature detection using overlapping convolution operators applied to static images do not closely resemble the human visual system.

Recent research has focussed on the use of hexagonal pixel-based images [7, 11] as the hexagonal pixel lattice closely resembles the structure of the human fovea and has many advantages in terms of image capture and analysis. In [10], Sheridan introduced a unique addressing system for hexagonal pixel based images, known as the spiral architecture, that addresses each hexagonal pixel with a single co-ordinate address, rather than the two co-ordinate address scheme typically used with rectangular image structures. Such a one-dimensional addressing scheme potentially provides an

appropriate structure for real-time image processing of hexagonal images. However, with respect to feature extraction via convolution, where typically an operator is applied to a pixel and its neighbours, the process of determining these neighbours in a one-dimensional addressing scheme is not always trivial and can require time consuming special hexagonal and radix-7 addition. In [9] the concept of eye tremor - rhythmic oscillations of the eye - has been exploited for image processing. Instead of applying the operators to every pixel in an overlapping manner typical of standard convolution, Roka et al. used nine overlapping images and applied the masks in a non-overlapping manner; however, they still assume the image to be comprised of rectangular pixels that are addressed in a two-dimensional structure, and they apply standard square image processing masks.

We present a biologically motivated approach to feature detection based on the use of the spiral architecture in conjunction with eye tremor and convolution of non-overlapping Laplacian masks [9]. The Laplacian operator can be considered analogous to the on-off receptive fields found in the retina in which the centre of the receptive field is negative, surrounded by positive values or vice versa. We develop a *cluster operator*, based on the spiral architecture, which can be applied to a one-dimensional spiral image in a fast and efficient way.

2 Spiral Architecture

In the spiral architecture [10] the addressing scheme originates at the centre of the hexagonal image (pixel index 0) and spirals out using one-dimensional indexing. Pixel 0 may be considered as a layer 0 tile. Pixel 0, together with its six immediate neighbours indexed in a clockwise direction (pixels 1, ..., 6) then form a layer 1 super-tile centred at pixel 0. This layer 1 super-tile may then be combined with its six immediately neighbouring layer 1 super-tiles, the centres of which are indexed as 10, 20, 30, 40, 50 and 60; the remaining pixels in each of these layer 1 super-tiles are then indexed in a clockwise direction in the same fashion as the layer 1 super-tile centred at 0, (e.g., for the layer 1 super-tile centred at 30, the pixel indices are 30, 31, 32, 33, 34, 35 and 36).

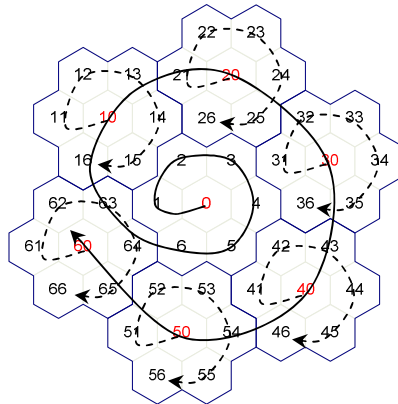


Fig. 1. One-dimensional addressing scheme in the central region of the image, showing one layer 2 super-tile, comprising 7 layer 1 super-tiles

The entire spiral addressing scheme is generated by recursive use of the super-tiles; for example, seven layer 2 super-tiles are combined to form a layer 3 super-tile. Ultimately the entire hexagonal image may be considered to be a layer L super-tile centred at 0 comprising 7^L pixels. Figure 1 shows the spiral addressing scheme for the central portion of an image (up to the layer 2 super-tile).

A major advantage of this addressing scheme is that any location in the image can be represented by a single co-ordinate value. This is advantageous for a number of reasons: it permits full exploitation of the symmetry of the hexagonal lattice; placement of the origin at the centre of the image simplifies geometric transformations such as rotation and translation; and most importantly it allows the spiral image to be stored as a vector [7]. Spatially neighbouring pixels within the 7-pixel layer 1 super-tiles in the image remain neighbouring pixels in the one-dimensional image storage vector. This is a very useful characteristic when performing image processing tasks on the stored image vector, and this contiguity property lies at the heart of our approach to achieve fast and efficient processing for feature extraction.

3 Cluster Operators

We refer to the operator that is applied to a cluster neighbourhood in the spiral architecture as a *cluster operator*. In recent work [2] we have shown how a finite element based approach can be used to create hexagonal operators based on constructing either two independent directional derivative operators aligned in the x- and y- directions, or tri-directional operators aligned along the x-, y- and z- hexagonal axes. Our operators are built using a regular mesh of equilateral triangles with nodes placed at the centres of each hexagonal pixel. With each node s we associate a piecewise linear basis function ϕ_s , with $\phi_s = 1$ at node s and $\phi_s = 0$ at all other nodes $t \neq s$. Each ϕ_s is thus a "tent-shaped" function with support restricted to a small neighbourhood of six triangular elements centred on node s . We represent the image by a function $I = \sum_{q \in Q} I(q)\phi_q$, where Q denotes the set of all nodal addresses; the parameters $\{I(q)\}$ are the image intensity values at the pixel centres.

Feature detection and enhancement operators are often based on first or second order derivative approximations, and we consider a weak form of the second order directional derivative $-\underline{\nabla} \cdot (\mathbf{B}\underline{\nabla}u)$ over small neighbourhoods. To approximate the second directional derivative, $-\underline{\nabla} \cdot (\mathbf{B}\underline{\nabla}u)$, over a λ -neighbourhood $N_\lambda(s)$ centred on the pixel with spiral address s , the respective derivative term is multiplied by a test function $v \in H^1$ and the result is integrated over $N_\lambda(s)$. The neighbourhood size λ corresponds to the layer λ : here we focus only on the neighbourhood $N_1(s)$ containing seven pixels (Figure 1 shows seven layer 1 neighbourhoods with $s = 0, 10, 20, 30, 40, 50$ and 60). Hence at each node s we may obtain a layer λ weak second order directional derivative $D_\lambda(s)$ as

$$D_\lambda(s) = \int_{N_\lambda(s)} (\mathbf{B}\nabla I) \cdot \nabla \psi_s^\lambda d\Omega \tag{1}$$

where $\mathbf{B} = \underline{b} \underline{b}^T$ and $\underline{b} = (\cos \theta, \sin \theta)$ is the unit direction vector. Our cluster operator is the isotropic form of the second order derivative, namely the Laplacian $-\nabla \cdot (\nabla u)$; this is equivalent to the general form in which the matrix \mathbf{B} is the identity matrix \mathbf{I} .

Although we are not addressing the issue of scale at this stage, it should be noted that each neighbourhood test function ψ_s^λ is restricted to have support over the neighbourhood $N_\lambda(s)$ for any choice of layer $\lambda = 1, 2, 3, \dots$. Thus we may write

$$D_\lambda(s) = \sum_{q \in Q} \left(I(q) \int_{N_\lambda(s)} (\nabla \phi_q) \cdot \nabla \psi_s^\lambda d\Omega \right) = \sum_{q \in N_\lambda(s)} H_\lambda(q) \times I(q) \tag{2}$$

where H_λ is a hexagonal operator of “size” λ (having the size and shape of a hexagonal λ -neighbourhood). Our general design procedure incorporates a layer-related parameter that enables scale to be addressed in future work. In this paper we have chosen each neighbourhood test function ψ_s^λ to be a Gaussian function centred on node s , parameterised so that 95% of its central cross section falls within $N_\lambda(s)$ and then its tails truncated so that support is restricted to the neighbourhood $N_\lambda(s)$. Figure 2 shows the layer 2 Laplacian cluster operator generated in this way.

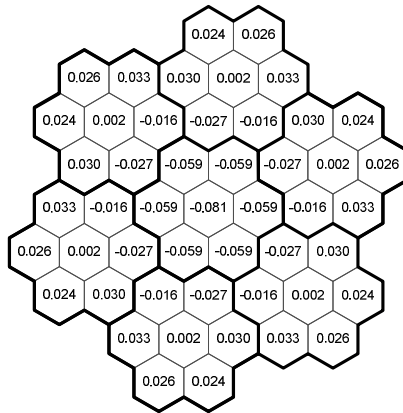


Fig. 2. Layer 2 Laplacian cluster operator

4 Framework for Fast Processing

Our proposed fast image processing framework has three essential elements: simulation of eye tremor by use of a set of slightly off-set images of the same scene; a special definition of sparse spiral convolution; and efficient identification of the spiral architecture pixel addresses within those $\lambda = 2$ neighbourhood clusters on which

local operator convolution is required to be performed. We consider each of these three elements in turn below.

4.1 Simulation of Eye Tremor

We consider the hexagonal image I_0 to be the “base” image, and we capture six further images, $I_j, j = 1, \dots, 6$, of the same scene. Each of these additional images is off-set spatially from I_0 by a distance of one pixel in the image plane along one of the three natural hexagonal axis directions. This mechanism simulates the phenomenon of “eye tremor”. In each image $I_j, j = 1, \dots, 6$, the pixel with spiral address 0 represents the same spatial location in the scene as the pixel with spiral address j in I_0 .

The centre (i.e., the pixel with spiral address zero) of each image $I_j, j = 0, \dots, 6$, is thus located at a pixel within the $layer \lambda = 1$ neighbourhood centred at the pixel with spiral address 0 in image I_0 , as shown in Figure 3.

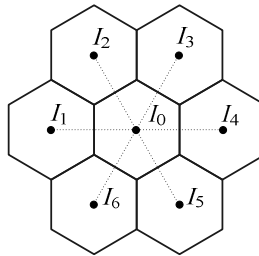


Fig. 3. The 7 image centres in the eye tremor approach

Through use of the spiral architecture for pixel addressing, it is assumed that image I_0 is stored in a one-dimensional vector (with base-7 indexing), as shown in Figure 4. Using the spiral architecture the additional images $I_j, j = 1, \dots, 6$, are stored similarly.

0	1	2	3	4	5	6	10	...	16	20	...	26	60	...	66	100	...	106	110	...
---	---	---	---	---	---	---	----	-----	----	----	-----	----	-----	-----	----	-----	----	-----	-----	-----	-----	-----

Fig. 4. One-dimensional storage vector showing address values for image I_0

4.2 Sparse Spiral Convolution

For a given image I_0 , convolution of a hexagonal operator H_λ of “size” λ (having the size and shape of a layer λ cluster neighbourhood) across the entire image plane is achieved by convolving the operator sparsely with each of the seven images $I_j, j = 0, \dots, 6$ and then combining the resultant outputs.

In each of the images $I_j, j = 0, \dots, 6$, we apply the operator H_λ only when centred at those pixels with spiral address 0 (mod 7). Figure 5 shows a sample of pixels in

image I_0 for which the label $j = 0, \dots, 6$ for each pixel indicates in which of the images $I_j, j = 0, \dots, 6$, the pixel address takes the value $0 \pmod{7}$. Each pixel in image I_0 may be thus uniquely labeled.

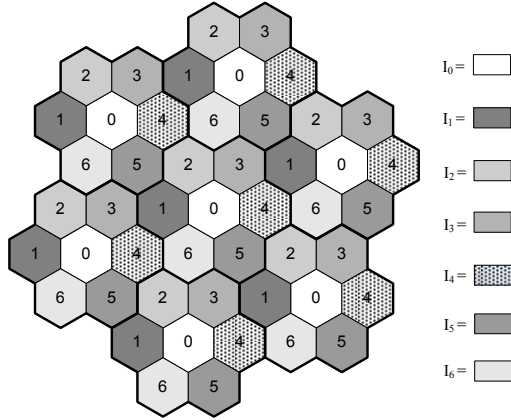


Fig. 5. Pixel positions in image I_0 corresponding to pixels in images $I_j, j = 0, \dots, 6$ with address value $0 \pmod{7}$.

We then define “base-7-zero” convolution

$$S_\lambda = H_\lambda \otimes_{\tau_0} I \tag{3}$$

of the hexagonal operator H_λ with the hexagonal image I by:

$$S_\lambda(s_0) = \sum_{s \in N_\lambda(s_0)} H_\lambda(s) \times I(s) \quad \forall s_0 \in \{s | s = 0 \pmod{7}\} \tag{4}$$

and $N_\lambda(s)$ denotes the layer λ neighbourhood cluster centred on the pixel with spiral address s in image I . Thus in order to implement “eye tremor” simulation in relation to the base image I_0 , we apply “base-7-zero” convolution (\otimes_{τ_0}) of the hexagonal operator H_λ with each of the hexagonal images $I_j, j = 0, \dots, 6$, thus generating seven output responses:

$$S_\lambda^j = H_\lambda \otimes_{\tau_0} I_j, j = 0, \dots, 6 \tag{5}$$

that are combined to provide the consolidated response $E_\lambda = H_\lambda \otimes I_0$.

4.3 Layer 2 Cluster Neighbourhood Address Identification

As it is not appropriate to construct a Laplacian operator as small as Layer 1 due to its sensitivity to noise [12], we focus on a layer 2 Laplacian cluster operator. Application of a layer $\lambda = 2$ operator on a neighbourhood $N_2(s_0)$ requires identification of an

ordered set of addresses of the centres of the layer 1 cluster neighbourhoods contained in $N_2(s_0)$. As $s_0 = 0 \pmod 7$ we may determine these 7 centres of layer 1 cluster neighbourhoods as:

$$c_\alpha = s_0 + 10i_\alpha, \quad i_\alpha = 0, \dots, 6 \tag{6}$$

For each layer 1 centre, c_α , the corresponding layer 0 cluster neighbourhood addresses are then simply given by the ordered set

$$\{c_\alpha + j\}_{j=0}^6 \tag{7}$$

From the above it can be seen that the amount of special hexagonal addition required to identify an ordered list of the pixel addresses in a layer 2 cluster neighbourhood $N_2(s_0)$ centred on a pixel with spiral architecture address s_0 ($s_0 = 0 \pmod 7$) is considerably less than would typically be required to identify such an address list for an arbitrary layer 2 cluster neighbourhood $N_2(s)$ with $s \neq 0 \pmod 7$. (For $s \neq 0 \pmod 7$, a full set of 49 special hexagonal additions would be required.) Hence, as demonstrated by the performance evaluation results presented in Section 6, our proposed approach using sparse “base-7-zero” convolution is significantly more efficient than standard spiral convolution.

5 Spiral Implementation

In terms of implementation using the one-dimensional vector structure for the images $I_j, j = 0, \dots, 6$, that is facilitated by the spiral architecture, each output response $S_\lambda^j, j = 0, \dots, 6$ is stored in a one-dimensional vector “a-trous” with non-empty values corresponding to the array positions with indices $0 \pmod 7$, as illustrated in Figure 6.

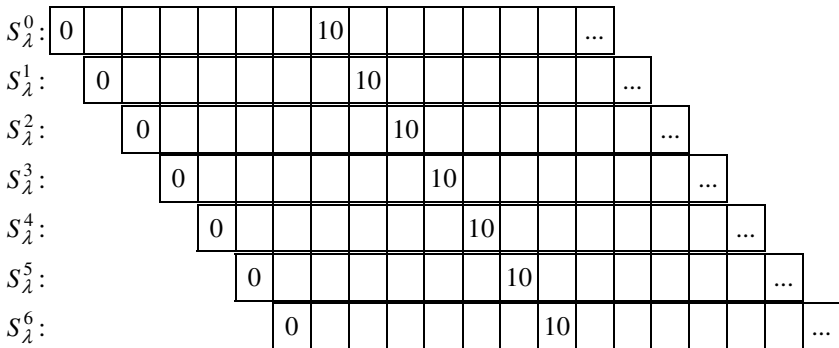


Fig. 6. Assembly of the one-dimensional vectors “a-trous” $S_\lambda^j, j = 0, \dots, 6$

The one-dimensional vectors “a-trous” $S_\lambda^j, j = 0, \dots, 6$, may then be assembled according to the “shifted” structure illustrated in Figure 6:

$$\forall s_0 \in \{s | s = 0(\text{mod } 7)\}, E_\lambda(s_0 + k) = S_\lambda^k(s_0) \text{ for } k = 0, \dots, 6 \tag{8}$$

to yield the consolidated output image $E_\lambda(I_0) = H_\lambda \otimes I_0$ as shown in Figure 7.

$S_\lambda^0(0)$	$S_\lambda^1(0)$	$S_\lambda^2(0)$	$S_\lambda^3(0)$	$S_\lambda^4(0)$	$S_\lambda^5(0)$	$S_\lambda^6(0)$	$S_\lambda^0(10)$	$S_\lambda^1(10)$
------------------	------------------	------------------	------------------	------------------	------------------	------------------	-------------------	-------------------	-----	-----

Fig. 7. Consolidated output image $E_\lambda(I_0) = H_\lambda \otimes I_0$ resulting from assembly of the vectors “*a-trous*” in Fig 6

6 Performance Evaluation

We present run-times for our proposed biologically motivated approach in comparison with standard convolution of an operator with a spiral image where the pixel neighbour addresses are found in two different ways: (i) via standard spiral convolution using special hexagonal and radix-7 addition; and (ii) neighbours are stored in a look-up table (LUT). The LUT takes 8.001s to generate for the spiral convolution approach and 0.091s for the eye tremor approach in the case of $\lambda = 2$, as the spiral convolution LUT requires all 49 addresses to be stored per record whereas the eye tremor LUT requires only the addresses of the centres of the seven $\lambda = 1$ sub-clusters. Algorithmic run-times were recorded for application of $\lambda=2$ Laplacian cluster operators to the *clock* image shown in Figure 8a. The hexagonal *clock* image was obtained by resampling the original square pixel-based image to a spiral image containing 117649 hexagonal pixels. The edge map and each of the seven “*a-trous*” gradient outputs used to generate it are also shown in Figure 8. The run-times presented in Table 1 are the averages over 100 runs using a workstation with a 2.99Ghz Pentium D processor and 3.50Gb of RAM.

Table 1. Algorithm run-times for 49-point operator ($\lambda = 2$)

Method	Run-time
Biologically motivated “eye tremor” approach	0.106s
Standard spiral convolution	25.087s
“Eye tremor” approach using LUT	0.028s
Spiral convolution using LUT	0.018s

The results in Table 1 demonstrate that our biologically motivated approach is approximately 250 times faster than standard spiral convolution. The implementation of both approaches can be accelerated by use of a LUT; this results in the run-time for the spiral convolution approach being faster than our approach but there is also, of course, the additional overhead of storing a larger LUT, which can be considerable for larger values of λ . More specifically though, the run-time of 0.028s for our approach is a combined time for processing seven “*a-trous*” gradient outputs; however, for video processing, as the spiral addressing in each frame in a sequence can be off-set

slightly from its adjacent frames in a cyclic pattern, once the first seven *a-trous* images are processed to generate a complete frame, the addition of each subsequent *a-trous* image will generate a new complete frame. Hence each subsequent frame will be generated in one seventh of the time stated in Table 1, i.e., approx 0.004s. Therefore, our biologically motivated approach will be approximately four times faster than full implementation of the spiral convolution LUT approach when processing a stream of video images.

Additionally, using an adaptation of the Figure of Merit [1], we compare the accuracy of the Layer 2 Laplacian cluster operator (denoted as L2) with two Laplacian operators designed for use on standard rectangular-based images: the 7×7 Marr Hildreth operator (denoted as MH7) and the 7×7 Laplacian near-circular operator [2] (denoted as LNC7), which are both equivalent in size to the Layer 2 operator (L2, MH7 and LNC7 all have 49 mask values). In order to measure accurately the performance of the Laplacian cluster operator, we have modified the

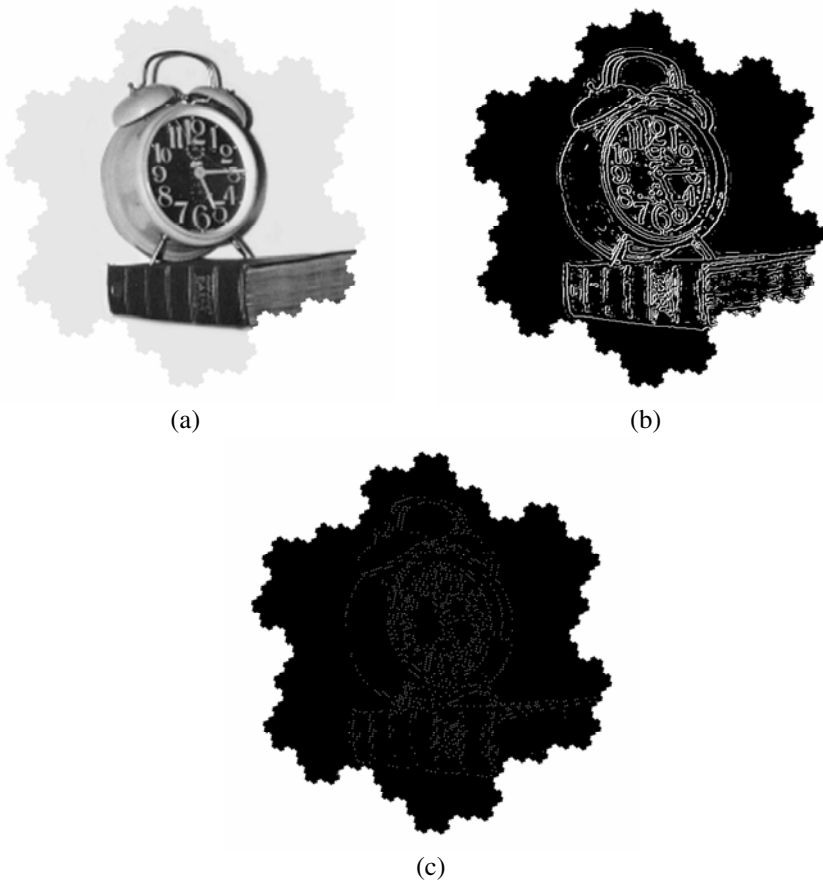


Fig. 8. (a) Original image; (b) Completed spiral edge map; (c) S_0 , an example of a corresponding edge map “*a-trous*”

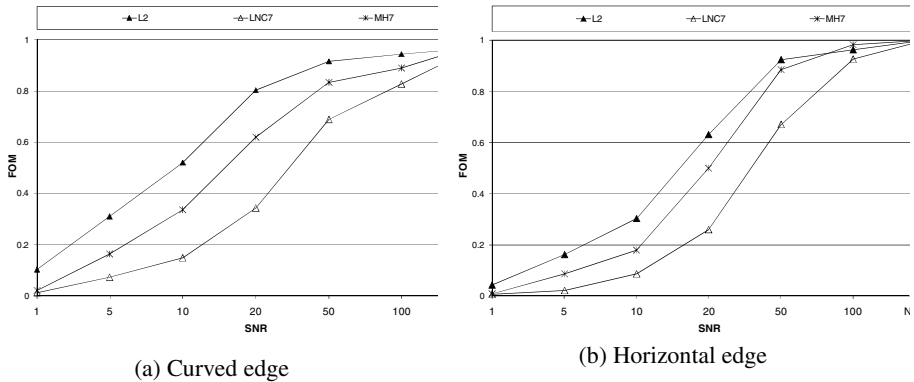


Fig. 9. Figure of Merit results comparing the tri-directional spiral operators (L2) with equivalent standard use of square operators (LNC7, MH7) using synthetic images containing (a) a curved edge; (b) a horizontal edge

well-known Figure of Merit technique in [1] to accommodate the use of hexagonal pixel-based images. The Figure of Merit results shown in Figure 9 illustrate that the proposed L2 spiral Laplacian operator has increased edge locational accuracy over the equivalent rectangular operators for all evaluated edge directions.

7 Conclusion

We present a biologically motivated approach to fast feature extraction using the concept of eye tremor. We have presented a design procedure for Laplacian cluster operators for use within our fast framework. The Figure of Merit results show that the Laplacian cluster operator provides better edge detection performance than the other square Laplacian masks of equivalent size. We have demonstrated that the approach of generating feature maps “*a-trous*”, using Layer 2 cluster operators, that can be combined into a single complete feature map is significantly faster than standard convolution or use of a neighbourhood LUT on hexagonal images. Generalisation of the approach to enable application of Laplacian operators at various scales will be the focus of future work.

References

- [1] Abdou, I., Pratt, W.: Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of the IEEE* 67(5), 753–763 (1979)
- [2] Coleman, S.A., Gardiner, B., Scotney, B.W.: Adaptive Tri-Direction Edge Detection Operators based on the Spiral Architecture. In: *IEEE ICIP*, pp. 141–144 (2010)
- [3] Coleman, S.A., Scotney, B.W., Herron, M.G.: A Systematic Design Procedure for Scalable Near-Circular Laplacian of Gaussian Operators. In: *ICPR*, pp. 700–703 (2004)
- [4] Curcio, C.A., et al.: Human Photoreceptor Topography. *Journal of Comparative Neurology* 292, 497–523 (1990)

- [5] Dacey, D.M., Packer, O.S.: Receptive Field Structure of h1 Horizontal Cells in Macaque Monkey Retina. *Journal of Vision* 2(4), 279–292 (2000)
- [6] Hirsch, J., Miller, W.H.: Does Cone Positional Disorder Limit Resolution? *Journal Optical Soc. of America A: Optics, Image Sci, and Vision* 4, 1481–1492 (1987)
- [7] Middleton, L., Sivaswamy, J.: *Hexagonal Image Processing; A Practical Approach*. Springer, Heidelberg (2005)
- [8] Mollon, J.D., Bowmaker, J.K.: The Spatial Arrangement of Cones in the Primate Fovea. *Nature* 360, 677–679 (1992)
- [9] Roka, A., et al.: Edge Detection Model Based on Involuntary Eye Movements of the Eye-Retina System. *Acta Polytechnica Hungarica* 4(1), 31–46 (2007)
- [10] Sheridan, P.: *Spiral Architecture for Machine Vision*. Ph. D. Thesis, University of Technology, Sydney (1996)
- [11] Shima, T., Saito, S., Nakajima, M.: Design and Evaluation of More Accurate Gradient Operators on Hexagonal Lattices. *IEEE Trans. PAMI* 32(6), 961–973 (2010)
- [12] Vernon, D.: *Machine Vision*. Prentice Hall International (UK) Ltd, Englewood Cliffs (1991)