An Adaptable Foveating Vision Chip

Timothy G. Constandinou, Patrick Degenaar and Chris Toumazou

The Institute of Biomedical Engineering
Imperial College of Science, Technology and Medicine
Exhibition Road, London SW7 2AZ, United Kingdom.
Email: {t.constandinou,p.degenaar,c.toumazou}@imperial.ac.uk

Abstract— In this work we present an adaptable foveating vision chip. This chip has no physical foveation; all its pixels are in the same uniform pattern. However with a given input signal it is possible to define areas of the chip which act as a fovea, returning high spatial resolution. The surrounding peripheral vision acts to return lower spatial resolution but much higher temporal resolution. Our chip is therefore able to achieve full spatial resolution via scanning of the fovea across the visual field. This operation is analogous to the functioning of the human eye. In the human eye however, the limitations of biology enforce a fixed fovea, while the optomechanics are highly efficient. In our structure, we acknowledge the limitations of physical optomechanics but use the advantages of silicon processing to achieve dynamic foveation. This paper discusses the algorithm, its implementation and simulated results describing its responses and power consumption.

I. INTRODUCTION

The recent dramatic rise in popularity of digital imaging has been the result of the quality and cost performance of charge coupled device (CCD) technology. CCD's have improved to be very capable for high-end imaging applications; for example in taking high quality colour images at mega-pixel resolutions. However, Complementary Metal Oxide Semiconductor (CMOS) imagers offer superior integration, power dissipation and system size at potentially lower cost. The integration of the phototransduction elements with image processing, in recent years, has prompted vigorous activity in the development of smart CMOS imagers. The neuromorphic vision community has been taking the process a step further by attempting to implement algorithms inspired by animal visual systems to develop intelligent vision chips.

The human eye is capable of detecting more than 8 orders of magnitude of light intensity. It can achieve this with frequency responses of 25Hz in perceptive vision and up to 150Hz in microsaccades. This remarkable specification is rooted in both the biological capabilities of the rhodopsin photocascade of the rods and cones and the way the retina processes information.

Foveation allows the advantage of trading spatial resolution for temporal resolution in a limited bandwidth regime. Previous work [1-5] has concentrated on mimicking biology, developing foveated vision chips by physically defining a foveal region of increased resolution at the centre of the CMOS imager. The

biological eye has exceptionally good optomechanics. This enables the eye to be moved round the visual field at high speeds. In contrast, even with modern MEMS technologies, it is challenging to position silicon-based cameras with as many degrees of freedom at such high speed and accuracy, in addition to being compact. In silicon however, it is possible to electronically define and reposition the fovea, something difficult in biology and not seen in nature. Thus, the authors of this paper believe that the best approach to artificial foveation is electronic rather than physical.
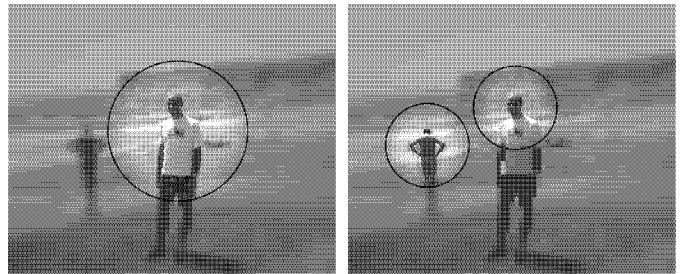

Fig. 1: The effect of foveation on imaging. Shown is: (on the left), imaging with a single large fovea and (on the right), imaging with two foveal regions.

Adaptable foveation allows for an area to be specified as either a high spatial resolution fovea or high temporal resolution periphery. The scheme developed in this paper acts not only to designate a spot which is to act as a fovea, but to be able to dynamically adjust the size of that region at will. The broadening or thinning of the fovea spot is adaptable to imaging requirements and requires only a simple coordinate in feedback.

The basic principle is to group pixels into photoreceptor clusters and to take intensity information on an asynchronous event timing basis rather than through raster scanning. The photoreceptor clusters can either let all their constituents send out spatial information or integrate their response into a single combined output. A global current spreading network between photoreceptor clusters can be used to determine which clusters relative to an initial position get defined as fovea or periphery.

II. SYSTEM ALGORITHM

Our scheme proposes to combine pixels into repeat unit structures of 3×3 pixels. The 9 pixels in these unit structures can act as either individual photoreceptors or as a singular

compound photo-sensor. When combined, the photocurrents are added for better light sensitivity, and the output can be 9 times faster at the cost of 9 times less spatial resolution. In this way the unit structures can be defined as foveal (individual sensors) or peripheral (singular sensors). This scheme is visualized in Fig. 2 below.
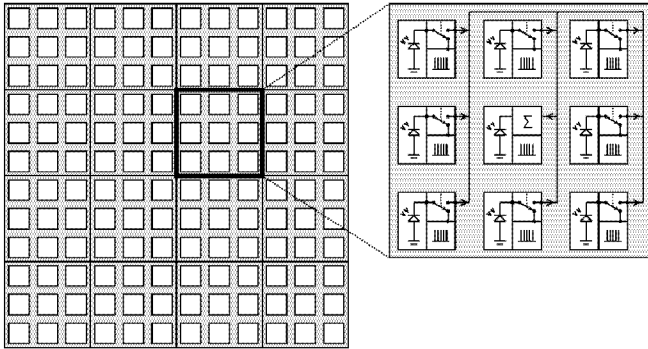

Fig. 2: The scheme of how the photoreceptor groups can be reconfigured to operate as either foveal or periphery units.

To achieve the foveal and peripheral definition, we use a current spreading and thresholding technique. Photoreceptor groups within the array can be addressed using (x,y) coordinates. On addressing a specific group, a foveal current is sourced to that group; subsequently being attenuated and spread to neighbouring photoreceptor groups. Thus there is a decay function from the designated fovea to the furthest periphery. Given that the fovea is not predetermined, each point will feed back to its surrounding photoreceptor groups as in Fig. 3 below. A thresholding function can then be used to differentiate between the fovea and the periphery.
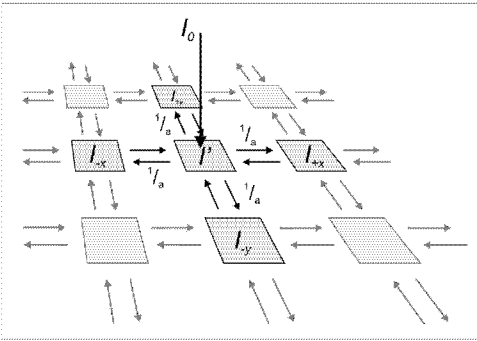

Fig. 3: Foveal current spreading technique.

Mathematically, the foveal decay current can be described in terms of a recursive spreading function, where each component's feed-forward influence from its preceding neighbour is:

$$I' = I_{+x+y} \sum_{n=0}^{\infty} \frac{(n+1)}{a^{(2n+1)}} + I_{-x+y} \sum_{n=0}^{\infty} \frac{(n+1)}{a^{(2n+1)}}$$
$$+ I_{+x-y} \sum_{n=0}^{\infty} \frac{(n+1)}{a^{(2n+1)}} + I_{-x-y} \sum_{n=0}^{\infty} \frac{(n+1)}{a^{(2n+1)}}$$

Where I_{xy} is interactive influence from one of the neighbouring points in the grid and a is the attenuation factor. For the array to be stable at the central point the attenuation factor must be greater than 4.

These function relationships can be solved using iterative recursion in standard mathematical software. Results of these simulations can be seen in Fig.8. For our purposes, we have used an initial attenuation factor of 4 at the fovea, followed by an attenuation factor of 6 between all other connections.
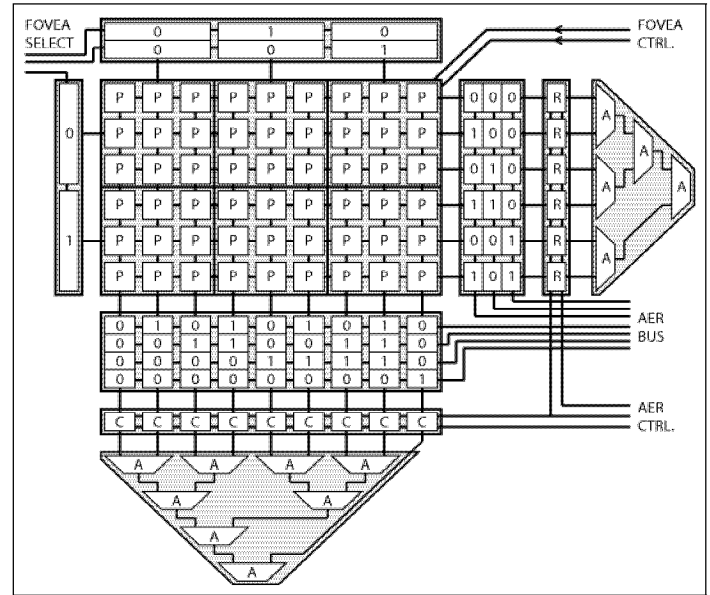

Fig. 4: System architecture of the adaptable foveating vision chip. Sub-blocks illustrated include: P=Pixel, R=Row Latches, C=Column Latches, A=Arbiters and 0/1=Address Encoders or Decoders.

III. CIRCUIT IMPLEMENTATION

A. Top level System Architecture

The address system for our chip is similar to that of previous [6-7] asynchronous Address Event Representation (AER) type imaging chips. When a pixel signals an event, a competition is relayed to the row and column arbitration trees; selecting a single row and column (in case of collision) that latches a single x and y plane header set, outputting the pulse x and y coordinates to the AER bus. This scheme is illustrated (architecturally) in Fig. 4. The main functional difference in this architecture from those previous is that here several address events are inhibited from signalling. This occurs when pixel groups act as peripheral pixels, sending their information as a single address event instead of using 9 distinct address events.

B. Foveation Control Circuit

The foveal control circuit, as previously explained, sources a current to a specific location spreading through a network of

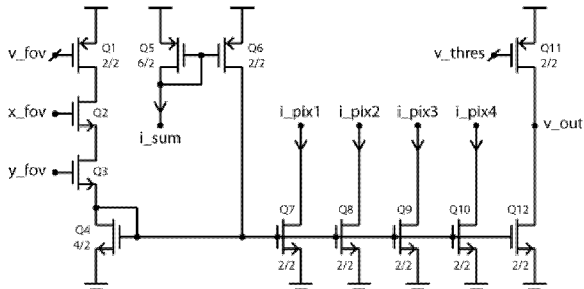interconnecting pixel groups. The circuit used to achieve this functionality is illustrated in Fig. 5.

Fig. 5: The implementation of the current spreading network determining the foveation state.

In this scheme, *v_fov* is used to set the seeding foveal current, defining the initial current level and therefore defining the fovea shape. The *x_fov* and *y_fov* inputs provide the row and column coordinates that act to select the foveal centroid. The foveal current is subsequently attenuated by a factor of 2 and distributed to each of the neighbouring pixels via simple mirror Q7–Q10. Furthermore, the received currents from adjacent cells also act to influence a cells effective foveal current contribution through the received *i_sum* which is attenuated by a factor of 6. For pixel groups situated away from the specified foveal centre, this influence provides the only source of current, as the initial foveal current is supplied to only a single pixel group. This effective current is then thresholded using a simple current comparator; implemented using an opposing CMOS pair Q11-Q12. This operates on the principle that the effective foveal current will be sunk from a current source Q11, set by *v_thres*, acting to drive either Q11 or Q12 into its ohmic region. This will therefore result in a voltage swing at node *v_out*, providing the foveal control signal.

This circuit is akin to previous pseudo-resistive networks [8-9] except that it is in current mode and has some advantages for this structure. Reducing the mismatch on this circuit is not critical as the specific shape of the fovea is not significant. The overhead this circuit therefore places on the fill factor is subsequently not great.

C. Photoreceptor Circuit

The photoreceptor scheme adopted uses the current-feedback technique proposed by Culurciello et al. [10]. This is illustrated below in Fig. 6. Each pixel has its own spiking circuit, however when the photoreceptor group is acting as a peripheral unit, all 9 photocurrents are switched into the central spiking circuit and the remaining spiking circuits are inhibited. Briefly, the photocurrent acts to discharge the capacitor, causing *v_photo* to fall, and thresholded by the Q6-Q7 inverter providing current feedback to provide positive feedback to increase speed and reduce power at the crossover point. Once a spike is induced, it dissociates the current path by switches Q2-Q3. The photoreceptor is reset by the AER acknowledge after off-chip transmission of the previous event. The spiking rate is therefore directly correlated to the

photocurrent and thus the light intensity. The inhibiting NOR gate serves to disconnect the AER handshake to inactive spike generators in pixel groups defined as periphery.
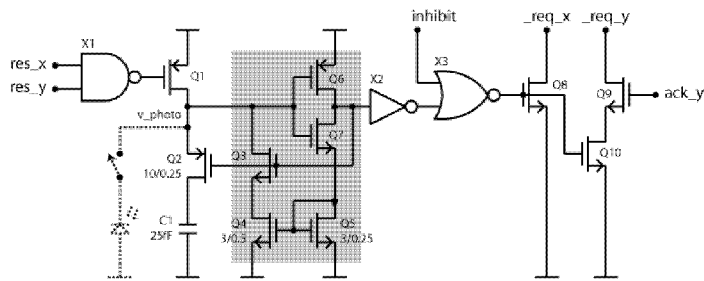

Fig. 6: The spiking photoreceptor circuit (shaded area illustrating the current-mode feedback [10]).

D. Layout

The surface fill factor is very important to the utility of this structure. We wish to achieve foveation in order to add utility, but not at the expense of scalability, for example to a mega-pixel resolution. To provide a proof of principle, a 27x48 pixel (9x16 photoreceptor cluster) array has been designed and submitted for fabrication in a standard 0.25μm CMOS process. The tessellating pixel-group layout is shown in Fig. 7 below.
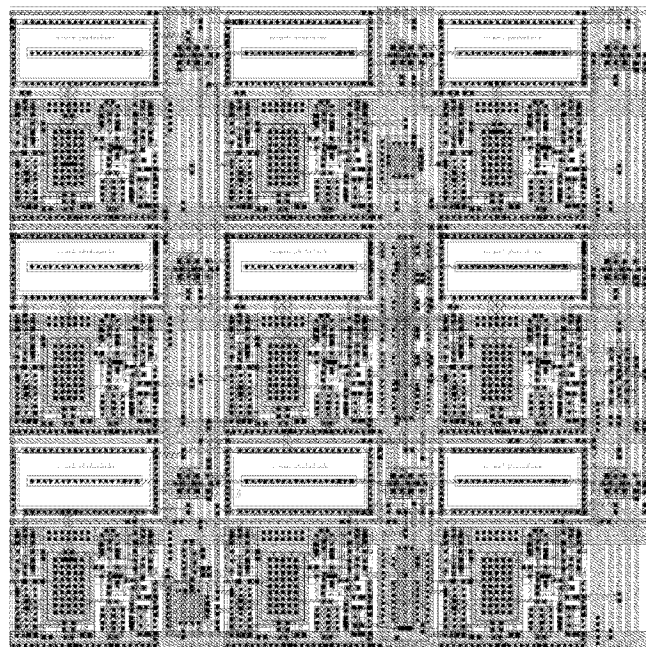

Fig. 7: The 3×3 pixel-group layout implemented in a 0.25μm CMOS process. The implemented surface fill factor of the photodiodes was 22.4%, but this is without major optimization. Fll factors of 30% or higher should be possible within the same overall cell size

Each individual photoreceptor within the 3×3 pixel-group has a dimension of 25×25 microns, including the foveation overhead. This compares favourably with commercial imaging array

3568

dimensions, allowing for implementation of a mega-pixel resolution at 25×25mm.

IV. SYSTEM SIMULATIONS

The foveal current spreading technique has been simulated for the 27×48 pixel array that has been sent for fabrication and also for a 320×200 pixel array. The simulation algorithm used a recursive iteration function to simulate both the forward and backward propagation of the foveal current. Results are illustrated in Fig. 8 below.
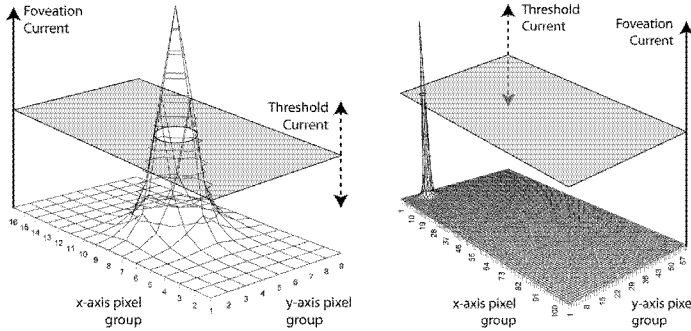

Fig. 8: The spreading function of the foveal current. Left, the spread for the 27x48 chip sent for fabrication, right, the spread for a 320x200 array

For the 27×48 pixel array the propagation forms a readily tuneable percentage of the array. However, this spreading function forms a small percentage of a 320×200 array. If desired the attenuation factor can be reduced to enhance spreading. In order to be fully scalable to a mega-pixel array, a logarithmic output function may be required to allow broadly tuneable spreading of the fovea.

V. DISCUSSION

In this paper we present a foveating chip which can adaptively change the size and position according to input stimuli. The output uses a standard AER imaging protocol and our total pixel sizes were 25×25 microns with a fill factor of 22.4%. The algorithm used for achieving foveation is simple, yet effective. Mismatch on the spread of the foveation scheme is not an issue, as a non-uniform fovea would not pose a problem.

We have shown that by foveating our chip, we can achieve areas specialized in either temporal or spatial resolution, listed below in the target design specification summary given in Table 1.

VI. CONCLUSION AND FURTHER RESEARCH

We have presented a biologically-inspired vision chip with the capability to adaptively foveate. We envision many application for our foveating technique to machine and neuromorphic vision applications. The ability to foveate allows the chip to have both high spatial and high temporal resolution within the framework of limited bandwidth.

We are presently looking at software algorithms to fully utilize our foveated system. We hope to devise intelligent feedback mechanisms to develop saccadic random walk type image construction. In addition our design allows the intriguing possibility of multiple foveation points. This could create "Alien Vision" opportunities that might be able to surpass the capabilities of the human retina.

Technology	UMC 0.25μm CMOS
Die size	2×1.5mm
Pixel array	48×27
Pixel size	25μm×25μm
Foveation (pixel group) array	16×9
Photodiode topology	deep n-well/p-substrate
Surface fill factor	22.4%
Spiking overhead	44.9%
Foveation & routing overhead	32.7%
Address-event bandwidth	$2×10^7$ events per second
Energy per spike	<1pJ
Foveation power overhead	<10% system consumption
High temporal region	8-bit @ 2kHz
High spatial region	8-bit @ 140Hz
High dynamic-range region	12-bit @ 5Hz

Table 1: Target design specifications for the adaptable foveating chip.

REFERENCES

[1] R. Wodnicki, G. W. Roberts and M. D. Levine, "A log-polar image sensor fabricated in a standard 1.2um ASIC CMOS process," IEEE Journal of Solid-State Circuits, vol. 32, no. 8, pp. 1274–1277, 1997.

[2] G. Sandini, J. Santos-Victor, T. Paidia and F. Berton, "OMNIVIEWS: direct omnidirectional imaging based on a retina-like sensor," Proceedings of IEEE Sensors, vol. 1, pp. 27–30, 2002.

[3] A. Bernardino, J. Santos-Victor and G. Sandini, "Foveated active tracking with redundant 2D motion parameters," Robotics and Autonomous Systems, vol. 3, no. 4, pp. 205–221, 2002.

[4] R. Etienne-Cummings, J. Van der Spiegel, P. Mueller and Z. Mao-Zhu, "A foveated silicon retina for 2D tracking," IEEE Transactions on Circuits and Systems II, vol. 47, no. 6, pp. 504–517, 2000.

[5] M. Azadmehr, J. P. Abrahamsen and P. Hafliger, "A Foveated AER Imager Chip," IEEE International Symposium on Circuits and Systems, pp. 2751–2754, 2005.

[6] P. Hafliger, "A Spike-based Learning Rule and its Implementation in Analog Hardware". PhD thesis, ETH Zurich, Switzerland, 2000.

[7] K. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," IEEE Transactions on Circuits and Systems II, Vol. 47, no. 5, pp. 416-434. 2000.

[8] A. Andreou and K. Boahen, "A 590 000 transistor 48 000 pixel, contrast sensitive, edge enhancing, CMOS imager-silicon retina," in Proc.16th Conf. Advanced Research in VLSI, pp. 225–240, 1995.

[9] E. A. Vittoz, "Pseudo-Resistive Networks and their Application to Analog Collective Computation," In Proc. of MicroNeuro, pp. 163-172, 1997.

[10] E. Culurciello, R. Etienne-Cummings, and K. Boahen, "A Biomorphic Digital Image Sensor," IEEE Journal of Solid State Circuits, vol 38, no 2, pp 281-294, 2003.