Challenges in the Evolution of Advanced Imaging Systems

Keith Lewis*
Sciovis Ltd., Victoria Road
Malvern, Worcestershire
WR14 2TE, UK
* Senior Fellow, QinetiQ, Malvern Technology Centre, Malvern WR14 3PS, UK
E-mail klewis@sciovis.com

ABSTRACT

Many of the more challenging goals set for future defence systems require a paradigm shift in imaging technology. Processes of bio-inspiration can inform the evolution of new imaging systems, especially those that exploit the benefits of computational imaging. Modern computing power shifts the emphasis away from costly highly engineered optical assemblies to lightweight systems exploiting algorithmic image reconstruction techniques. Yet some spiders are able to process several optical fields of different angular dimensions at the same time, which is a pre-requisite when organisms sense their environments using compound eye architectures. Some of the benefits of exploiting the parallels existing between evolutionary processes in biology and optical engineering are highlighted.

1. INTRODUCTION

Evolving military requirements are driving the need for imaging systems with enhanced levels of capability. For example, goals are being set for a compact lightweight and low cost airborne system with the capability of detecting and tracking many thousands of moving targets within a large urban environment. It must also provide sufficient discrimination to ensure that adequate track fidelity is maintained over extended periods of time. Such goals require a paradigm shift in imaging technology, which opens the door for exploring new ideas to guide the evolution of the imaging system. Biological vision systems have evolved over thousands of years, driven by the need to ensure the survival of the species. This paper explores aspects of that evolutionary process to support the development of new concepts for imaging systems.

2. EVOLUTIONARY PROCESSES IN BIOLOGY

The fossil records show that only three animal categories distinguishable by their external shapes existed on Earth 543 million years ago. Over a period of 43 million years, 35 additional categories emerged in what has been termed as the Cambrian Explosion. Significantly, the emergence of structural colour (eg via diffraction gratings) occurred at the same time as the emergence of vision. Some of the structures that give rise to colour persist to the present day. For example, the origin of iridescence in the sea-mouse is due to a regular array of hexagonally close-packed elements, almost identical to structures currently being “re-discovered” by science as photonic crystals [1]. Periodic arrays of melanin granules in the feathers of the peacock are also remarkably similar to photonic crystal log-pile structures.

As vision systems emerged, most organisms developed relatively simple sensors, initially only able to distinguish between light and dark. As time passed, more complex systems evolved. For example the eye of the nautilus [2] is based on a pinhole design, but with a curved multi-element sensing array that enables a real image to be produced, albeit subject to aberration. Camera-type eyes also emerged including single lens designs as in fish and birds, and the compound eyes of insects comprising hundreds of ommatidia. In the latter, the preferential alignment of rhodopsin within individual ommatidia forms the light detecting waveguide. The honeybee has over 5000 ommatidia [3], half of which are twisted 180˚ clockwise and half 180˚ anticlockwise. Each ommatidia has three visual cells for each colour (green, blue, ultra-violet), but one of the UV ommatidia is twisted by 40˚ and enables polarisation discrimination.

Adaptive Coded Aperture Imaging and Non-Imaging Sensors, edited by David P. Casasent, Timothy Clark,
Proc. of SPIE Vol. 6714, 671402, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.741457

The capability for polarisation discrimination also exists in fish, probably to enable clear vision through the water-air interface when looking for insects. In other cases, such as in the cuttlefish, the vision function is coupled to the complex display of patterns on the skin (containing both spectral and polarising effects), and a cuttlefish will retreat from its reflection in a mirror unless the degree of polarisation has been distorted by a suitable filter [4, 5].

Some sensing systems exploit distributed apertures to enable connectivity with other organs. An example here is the brittle star [6], whose surface is covered with an array of microlenses that focus light onto a bundle of nerves controlling pigment cells that migrate and darken the surface of the skin. In comparison, giant clams have several hundred small pinhole-type eyes on the exposed mantle, each with an angular acceptance angle of about 15˚ [7].

The resolving power of the human eye is largely a result of the high density of receptors on the retina. However it is the effectiveness of the eye/brain interface that determines the actual resolution achievable through the way that it manages connectivity to the individual receptors. The retinal neurons interconnecting the photoreceptors are organized in a number of layers, each processing the outputs of the rods and cones in a different way to provide information on brightness, edge detail, colour etc. The outputs are read-out by a layer of bipolar cells, which in turn feed directly into the retinal ganglion cells. There are some cross-links at this level, provided by groups of horizontal and amacrine cells. The horizontal cells perform a spatial smoothing of the inputs, whilst the bipolar cells act as difference amplifiers, using inputs from both the receptors and the horizontal cells, so providing a spatio-temporal filter. The ganglion cells are located at the front of the retina and provide an output to the optic nerve. On average there are 125 fewer ganglions than there are receptors, highlighting the importance of the cross-linking cells. In the foveal region the connectivity between receptor and ganglion is one-to-one, but the degree of convergence reduces considerably towards the edge of the retina. Foveal and non-foveal information is processed in different parts of the brain, with clear differences in latency.

At the highest level of performance, the human eye still falls short of that achieved by some raptors, which feature a receptor density of about 10^6/mm in the foveal region, more than five times that of the human being. The retina of the eagle has two foveal regions, and is flatter and larger that the human's, providing a much larger image [8]. The eagle can also change the shape of its cornea and lens, allowing it to be more precise in focusing and accommodation than the human. The spherical depressions in the foveae act like the convex lens in a Galilean telescope. The focal length of the birds' optical system can in theory exceed the axial length of the eye, thus providing the relatively large images necessary for high resolving power in a localised region of the retina. However even though an eagle can resolve a mouse at a distance of about 1km, it still relies on an oblique approach path to provide an enhanced tracking ability. Falcon eyes also have a high resolving power, with an ability to allow dynamic correction while diving at 200 mph. This again indicates something more about the role of the eye-brain interface and the extent to which visual information is being processed to enable the raptor to reach its prey.

The biological systems demonstrate that no single architecture dominates as a result of the evolution process and indeed there are many pathways that could have been pursued as organisms adapted themselves to their environments. However it is clear that forming an image is only part of the story and that biological vision systems exploit a considerable amount of information processing before the host organism is presented with the required picture of its environment. Even in the case of the smallest spider, the brain is able to process several optical fields of different angular dimensions at the same time. However for surveillance applications, any engineered system has to achieve a performance level comparable with that of existing systems, at least in terms of their functionality. For example, the VNIR narrow field of view channel in the Goodrich DB110 system has a focal length of 110", which enables it to resolve road vehicles at a distance of 70 nautical miles from a height of 40,000 feet [9]. But there is still a question about whether such levels of resolution are strictly required to enable vehicles to be tracked in an urban environment. There will always be a trade-off between spatial and temporal resolution as well as the latency associated with processing speed that determines the actual performance achievable. Indeed it was found in a study by Zhang et al [10] that the benefits of reducing resolution cell dimension below the standard deviation of the point spread function were minimal and that when the target intensity is low a larger pixel size is required to achieve the optimum probability of detection.

3. ENGINEERING OF COMPOUND EYE SYSTEMS – ROLE OF COMPUTATIONAL IMAGING

Many of the attributes of the biological compound eye have been explored in practical systems, such as the microlens architectures of Tanida et al [11] in their TOMBO system. Here the authors demonstrated colour imaging using a hybrid optical/digital system that exploited an array of microlenses to form a set of compound images on the detector, which could be suitably combined by post-detector processing. The working distance of the camera was determined by the focal length of the microlenses used (1.3mm). However care had to be taken to prevent interaction of the individual components of the compound image by the use of baffles. The output of the focal plane required careful post-processing to enable the individual images produced by the lenslets to be combined in such a way as to produce a high-resolution representation of the scene.

Jang and Javidi [12] have explored a variant of the TOMBO system. Here the authors have suggested using electronically synthesised lens arrays (ELSAs) to overcome Nyquist sampling limitations in a pixellated detector. This exploits earlier work where they proposed using a moving lenslet array to improve the quality of reconstructed 3D images. The ELSAs were generated as an array of Fresnel zone plates using a liquid crystal spatial light modulator (SLM). Each lenslet was produced by addressing of the appropriate pixels in the 1024 x 768 element SLM under computer control.

These techniques form a starting point for focal plane coding, which provides a structured pixel sampling process to enable non-degenerate sampling in multiple aperture systems. The group at Duke University [13] has been exploring such techniques and has also demonstrated the degree of resolution improvement that can be afforded through the use of sub-pixel contact masks.

An alternate approach follows the path set by the pinhole camera, which enables images free from chromatic aberration to be formed at all distances away from the pinhole, allowing the prospect of more compact imaging systems, with a much larger depth of field. However, the major penalty is the poor intensity throughput, which results from the small light gathering characteristics of the pinhole. Nevertheless, the camera is still able to produce images with a resolution determined by the diameter of the pinhole, although diffraction effects have to be considered. The light throughput of the system can be increased by several orders of magnitude while preserving angular resolution by using an array of pinholes. Each detector element sees the result of the summation of contributions from the various pinholes, corresponding to each viewpoint of the scene.

Early workers in the field were able to produce real images with the aid of digital de-convolution techniques. The concept of random aperture arrays was introduced in 1968 [14] and first exploited the use of digital encoding techniques. Other schemes included the non-redundant pinhole array [15], uniformly redundant arrays [16] and later non-redundant aperture techniques [17]. Unfortunately, practical problems associated with such hybrid imaging systems have restricted such systems in the past to X-ray and γ-ray applications, where there are few alternatives. The technology requires that every scene position be encoded on the detector in a unique way, placing constraints on the coded mask pattern and the manner in which its projections were detected. For optimal performance, the auto-correlation function of the mask pattern generally has a single peak with flat side lobes [18].

An example of a real system exploiting coded aperture techniques is the NASA Swift satellite Burst Array Telescope (BAT) camera. Launched in November 2004, the system uses a fixed coding aperture with 50,000 mask elements to locate cosmological γ-ray bursts. The mask is very large (ca 1.2 x 2.4 m) and is spaced from the 32,000-pixel detector array by a distance of 1m. The system has a 4.9 mrad point spread function but can locate γ-ray bursts to an accuracy of 0.3-1.2 mrad, so achieving significant subpixel localization [19].

It has been shown that high sensitivities can be achieved by scanning an array of wide-field coded aperture telescopes with aperture mask holes radially aligned to minimise auto-collimation effects [20]. Even greater improvement in sensitivity can be achieved by continuously scanning the field of view of the telescope over the source region to be imaged. Close to Poisson limited sensitivities can be achieved compared with limits of up to 100 times higher when dithering techniques alone are used. Results of this magnitude indicate the benefit that can be accrued from using adaptively coded aperture systems, provided that the technology is available at the component level to enable such architectures to be produced. There are also significant implications in relation to the amount of computing power needed for image reconstruction, since separate FFT operations may have to be carried out for each increment in field or mask shift.

The challenges associated with coded apertures at gamma ray energies are very different to the case when such architectures are proposed for visible or infrared radiation. In the former case, the masks are simple shadow devices, which impose geometric coding on the system, whereas in the latter case the role of diffraction becomes dominant, very much as in the case of the Young's slit experiment. However, although the diffraction removes the geometric mask coding, the coherent interactions between the multiple apertures in the mask result in fine structure associated with diffraction from the mask being present i.e. diffractive mask coding (figure 1). Phase information is encoded in the displacements of the Young's fringes so allowing an image to be obtained by reconstruction without having to resort to a lens.

For a given mask geometry, it is possible in principle to solve the diffraction integrals to calculate the intensity pattern incident on a detector array positioned behind the mask. For monochromatic radiation, the spatial structure can be characterized by an autocorrelation function and the angular width of the correlation peak is close to the diffraction limit. For broadband radiation, the angular width of the correlation peak increases by an amount inversely proportional to the square root of the distance z separating the mask and the detector array. The amount of power at high spatial frequencies is also reduced.

There are parallels between the Young's slit experiment and electron interferometry as exploited in the transmission electron microscope where it has been recognised that an out of focus image is an in-line hologram and that by exploiting this effect it has been possible to approach the resolution limit set by electronic instabilities of 0.1nm [21]. This hologram encodes the phase information through the use of a reference wave and enables an image to be produced by Fourier transform from far-field scattered amplitudes.

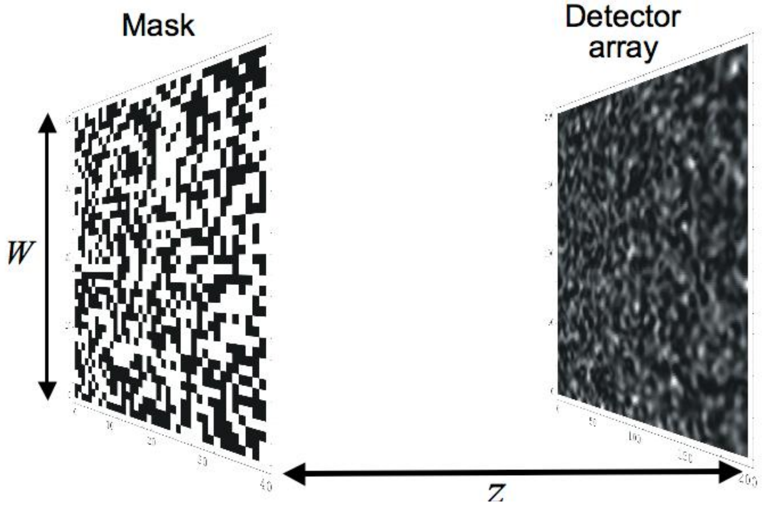

Figure 1 Schematic diagram highlighting propagation of radiation from mask to detector array in a diffractive coded aperture system

Spence et al have [22] explored the use of phase retrieval techniques in the optical and electron imaging regime using an iterative search algorithm - the hybrid input-output (HIO) algorithm of Gerchberg and Saxton [23]. Using a resolution test chart, illuminated by a collimated He-Ne laser, Spence et al were able to focus the diffracted light transmitted by the chart onto a CCD camera and reconstruct a digital image using the HIO algorithm. The algorithm recovered the phases of the object exit wave and always converged after about 50 iterations with different initial random phase sets. In these experiments a hard aperture was used to produce the reference constraint used in the reconstruction process. Images were also successfully reconstructed for greyscale objects. Using a field-emission electron microscope, the authors were also able to reconstruct an image of 20nm diameter apertures in a mask fabricated by electron beam lithography.

A challenge encountered in diffractive coded aperture systems is related to the compensation of the degree of optical loss in the system. Not only do losses occur as a result of the mark-space ratio of the mask design, but also because of the diffraction process itself, which can result in radiation being scattered outside the field of capture of the detector array. Thus a trade-off has to be made between the need to increase the value of z to achieve a given point spread function and the need to avoid too much loss of higher order diffractive components. There are also implications in relation to the signal to noise characteristics of the detector array. In a conventional imaging system, the refractive components provide significant optical gain (of the order of $10^5 – 10^6$), which concentrates light from a point source in the far field onto a single detector element. For the coded aperture case, the incident optical field from that distant point source is dispersed across the entire array. In the absence of detector noise this wouldn't be a problem for a 1 Mpixel array, but the effect of noise introduces a significant pedestal on the signal.

A key characteristic of the phase retrieval experiments of Spence et al [22] was the role hard objects of defined geometry to constrain the HIO algorithm. For the diffractive coded aperture system it is also necessary to provide a reference hologram to ensure that the system is properly calibrated. This can be done using a process similar to that used in astronomical telescopes where a laser guide-star forms the driver for wavefront reversal using phase conjugation techniques. However in a stable coded aperture system, this calibration process only has to be carried out once for each given mask design, unlike in the case of astronomy, where the process has to be repeated every 10 msec or so to compensate for the effects of atmospheric turbulence. There are some parallels here with the learning process in biological where the brain and its interface is trained to recognize its environment and all the cues so presented.

The discussion so far has been largely limited to the case of fixed masks. By varying the aperture pattern, other benefits can be realized, which open the pathway for improved imaging quality, variable field of view and steerable field of regard [24]. In a sense, some of these follow-on from the initial suggestion of Grindley and Hong [20], who found that up to two orders of magnitude improvement in signal to background ratio could be obtained in X-ray imaging by scanning the field of view over the telescope. A key enabler is the technology required for creating the adaptive mask and the degree of agility provided by that technology, not to say the increase in computational work-load required for the resulting imagery to be provided on a quasi-real time basis. In another paper at this conference, the group at QinetiQ [25] has explored the use of a transmissive MEMS-based shutter array, which has been configured for use in the mid-infrared. For the visible, it would be possible to consider reflective spatial light modulators based on liquid crystal on silicon technology or the Texas Instruments digital light projection engine.

4. BENEFITS OF CODED APERTURE IMAGING

Provided that some of the difficulties highlighted above can be overcome, there are significant benefits to be gained from the exploitation of coded aperture techniques in camera systems. The first arises because of the relative simplicity of the optical system itself, which removes much of the cost, weight and volume characteristics of a conventional lensed system. Modern processors are able to address the computational aspects of the technology and can even provide for real-time imaging capability. For the future, it is expected that even faster processors will become readily available, so paving the way for the exploitation of the coded aperture systems on light-weight unmanned vehicles. Another significant benefit is that it may not be necessary to process the output of the detector array on the platform itself, but that the raw data stream can be transmitted to a ground station for real-time processing, so ensuring the security of the imagery. The data stream is naturally encrypted and without knowledge of the geometry of the mask array, it would be extremely difficult to compromise the signal, especially when an adaptive mask is employed within the system.

Other benefits include:

- Infinite depth of field, 3-dimensional imaging of real world (low and mixed coherence) scenes – coded aperture systems will not require active illumination and can in principle exploit tomographic techniques to give 3D imaging.

- The lack of focusing means that the system is highly resistant to damage arising from pulsed lasers in the far field.

- Tolerance to point and line defects in the detector array, relaxing manufacturing tolerances with potential impact on system cost.

- Adaptive imaging allowing (for example) targeted priority scene acquisition (higher resolution, faster and/or at reduced system power and processing loads). Reconfigurable coded aperture configurations along with random access and windowing in CMOS arrays and photon tagging algorithms should allow adaptive prioritization of system bandwidth and performance and foveal-like functionality.

- A truly hybrid system approach where the same principle can be tuned/optimized for a variety of military applications including human visual and machine interpretation.

- A route to robust, conformal imaging systems – low system thickness, robustness and low rotational inertia will allow tiling of a variety of surfaces (e.g. helmet mounted imaging). Overlapping fields of view from adjacent tiles will allow additional routes to 3D imaging by triangulation.

5. CONCLUSIONS

Biological vision systems can be used to gain a useful insight when considering the development of new camera systems to meet demanding requirements. Modern computing power shifts the emphasis away from costly highly engineered optical assemblies to lightweight systems exploiting algorithmic image reconstruction techniques. The potential for lensless imaging is significant, both in meeting demanding defence requirements, but also for commercial applications such as those encountered in the provision of machine vision.

6. ACKNOWLEDGEMENTS

The author is indebted to Drs Chris Slinger, Philip Haskell, Jonathan Pritchard, Kevin Ridley, Neil Gordon, Geoff de Villiers, Malcolm Strens, Rebecca Wilson, Doug Payne, Mark McNie and other colleagues at QinetiQ, Tim Clark and Morley Stone at DARPA, and Dr Stanley Rogers (USAF, AFRL Sensors Directorate) for many helpful discussions.

7. REFERENCES

[1] A R Parker, R C McPhedran, D R McKenzie, L C Botten and N A Nicorovici: Nature 409 (6816) 36-37 (2001)

[2] D E Nilsson: Biosciences, 39 No. 5, (1989)

[3] K H Skrzipek, H Skrzipek: Z Zellforsch Mikrosk Anat 147 589-593 (1974)

[4] R T Hanlon and J B Messenger: Phil Trans Zool Soc Lond A **110** 17–35 (1988)

[5] N Shashar, P S Rutledge and T W Cronin: J Exp Biology 199 2077-2084 (1996)

[6] J Aizenberg, A Tkachenko, S Weiner, L Addadi, G Hendler: Nature. 412 819-822 (2001)

[7] M F Land: Proc Biol Sci. 270 (1511) 185–188 (2003)

[8] A W Snyder and W H Miller: Nature 275, 127 - 129 (1978)

[9] D Lange, M Iyengar, L Maver, G Dyer and J Francis: SPIE Proc 6546 654607 (2007)

[10] X Zhang, P Willett and Yaakov Bar-Shalom: SPIE Proc 5428 78 (2004)

[11] J Tanida, R Shogenji, Y Kitimara, K Yamada, M Miyamoto and S Miyatake: Optics Express 11 2109 (2003)

[12] J-S Jang and B Javidi: Optics Letters 27 1767 (2002)

[13] D J. Brady, M Feldman, N Pitsianis, J. P. Guo, A Portnoy and M Fiddy: Proc. SPIE 5907 590708 (2005)

[14] R H Dicke: Astrophys J 153, L101 (1968)

[15] M J E Goley: J Opt Soc Amer 61 272 (1971)

[16] T M Cannon and E E Fenimore: Opt Eng 19 283 (1980)

[17] K S Han, G J Berzins, D S Mason and D G Langner: Appl Optics 16 1260 (1977)

[18] E E Fenimore and T M Cannon: Applied Optics 17 337 (1978)

[19] D Palmer: Proc SPIE Conf 5199A-19, August 2003

[20] J E Grindley and J Hong Proc SPIE Conf 5168-43, August 2003

[21] A Orchowski, W D Rau and H Lichte: Phys Rev Lett 74 399-402 (1995)

[22] J C H Spence, U Weierstall and M. Howells: Phil Trans Royal Soc Lond A 360 875–895 (2002)

[23] R Gerchberg and W Saxton: Optik 35, 237 (1972)

[24] C W Slinger, N Gordon, K Lewis, G McDonald, M McNie, D Payne, K Ridley, M sterns, G de Villiers, R Wilson, and M Eismann: Proc SPIE 6714 6714007 (2007)

[25] M E McNie, D Combes, G Smith, N Price, K Brunson, K L Lewis, C W Slinger and S Rogers: proc SPIE 6714 6712012 (2007)