

# Optics Letters

## Multi-aperture foveated imaging

GUILLEM CARLES,<sup>1</sup> SHOUQIAN CHEN,<sup>1,2</sup> NICHOLAS BUSTIN,<sup>3</sup> JAMES DOWNING,<sup>4</sup> DUNCAN MCCALL,<sup>1</sup> ANDREW WOOD,<sup>3</sup> AND ANDREW R. HARVEY<sup>1,\*</sup>

<sup>1</sup>School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK

<sup>2</sup>School of Astronautics, Harbin Institute of Technology, Harbin 150001, China

<sup>3</sup>Qioptiq Ltd., St. Asaph LL17 0LL, UK

<sup>4</sup>ST Microelectronics, Edinburgh EH12 7BF, UK

\*Corresponding author: Andy. Harvey@glasgow.ac.uk

Received 9 February 2016; revised 3 March 2016; accepted 4 March 2016; posted 7 March 2016 (Doc. ID 259266); published 15 April 2016

**Foveated imaging, such as that evolved by biological systems to provide high angular resolution with a reduced space-bandwidth product, also offers advantages for man-made task-specific imaging. Foveated imaging systems using exclusively optical distortion are complex, bulky, and high cost, however. We demonstrate foveated imaging using a planar array of identical cameras combined with a prism array and superresolution reconstruction of a mosaicked image with a foveal variation in angular resolution of 5.9:1 and a quadrupling of the field of view. The combination of low-cost, mass-produced cameras and optics with computational image recovery offers enhanced capability of achieving large foveal ratios from compact, low-cost imaging systems.** © 2016 Optical Society of America

**OCIS codes:** (100.3020) Image reconstruction-restoration; (100.6640) Superresolution; (110.1758) Computational imaging.

<http://dx.doi.org/10.1364/OL.41.001869>

Conventional approaches to imaging typically aim for an approximately uniform spatial sampling frequency across the field of view, but for many applications, such as targeting, the salient requirement is for high-resolution imaging within a central, so-called foveal region of the image combined with a low-resolution periphery providing situational awareness and context. Foveated imaging offers more efficient use of a limited number of detector pixels or can be implemented as an image processing technique applied to conventional images to improve the efficiency of information transmission. Here our emphasis is to attain a large ratio between the spatial sampling frequency and image acuity for the central field of view (FOV) and a reduced sampling frequency at larger field angles. This mimics biological systems, such as the human visual system, where foveated imaging is associated with a variation in photoreceptor packing density that approximately mirrors the angular variation in the optical resolution of the eye.

Imaging systems with a variation in magnification of up to a factor of 2 between a central FOV and the periphery (the foveal ratio) have been demonstrated using conventional optical approaches, but higher ratios require dramatic increases in optical complexity. Higher foveal ratios are attractive for a wide range

of applications [1–6], and previous approaches include the use of multiresolution systems using single [1] or multiple sensors [2–4] and applications in microscopy [5,6]. In this Letter we report an experimental demonstration of computational construction of a foveated image using a multicamera array. Two mechanisms contribute to the high foveal ratio of 5.9:1 between the angular sampling frequency in the foveal and peripheral regions of the image: image distortion introduced by an array of prisms located in front of the camera array introduces nonuniform angular sampling by the sensor, and overlap of the field of regard of the cameras at the central FOV enables digital superresolution to increase the angular sampling rate at foveal regions. The use of mass-produced cameras and simple prisms enables high-performance foveated imaging at minimal cost.

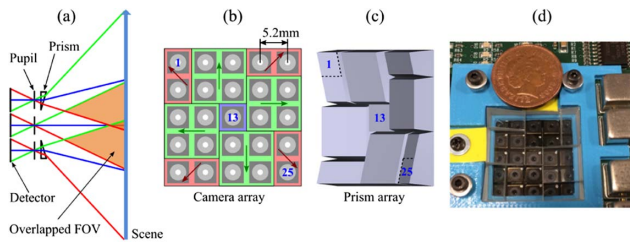
A  $5 \times 5$  multicamera array is assembled on a single printed circuit board, with the relatively low precision typical of an electronic assembly. An array of prisms is located in front of the camera array to angularly displace the field of regard observed by each individual camera. We have previously employed this camera array to demonstrate superresolution imaging using co-aligned cameras [7], but here the refraction by the prisms serves two functions: (1) by varying the prism angle from camera to camera, each camera can image a different field of regard so that an extended FOV is imaged as a mosaic from which a single extended-FOV image can be constructed; and (2) the refraction by the prisms introduces an optical distortion that contributes to the foveal characteristic of the image. Figure 1(a) shows a paraxial representation of the system operation.

The relationship between the semi-FOV of each camera,  $\phi_O$ , and the semi-FOV for the scene imaged through the adjacent prism,  $\phi_I$ , is

$$\sin(\phi_I + \theta) = n \sin \left( \arcsin \left( \frac{\sin(\phi_O)}{n} \right) + \theta \right), \quad (1)$$

where  $\theta$  is the prism angle and  $n$  is the refractive index of the prism. The optical distortions introduced by the prism are incorporated into the image registration used to construct the foveated, mosaicked image.

The arrays of  $5 \times 5$  cameras and prisms used here are depicted in Figs. 1(b) and 1(c). Several designs of prism array are possible: for example, 25 distinct prism angles could be used

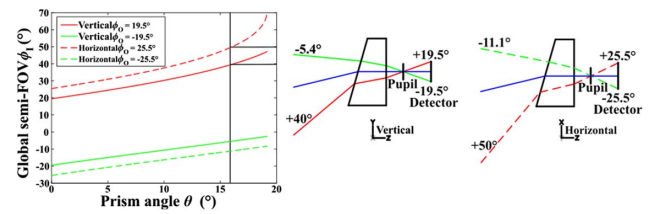


**Fig. 1.** (a) Principle of multi-aperture foveated imaging approach, (b), (c) layout of the camera and prism arrays, respectively, and (d) photograph of the assembled device. Only two types of prisms are used, highlighted by green and red in (b) and shown in (c); arrows in (b) indicate the direction of each prism slope.

to produce a mosaic of 25 images with a wide field of view. We chose here to combine nine subarrays of one, two, and four cameras, as shown in Figs. 1(b) and 1(c), and employ super-resolution in the overlapping fields of regard of multiple cameras. This simplifies the design of the prism array while proving a good foveal characteristic. As is described below, this results in varying degrees of overlap throughout the FOV, and consequently superresolution also varies the image spatial sampling frequency in a foveated fashion, in addition to the foveation provided by optical distortion. In this case two types of prisms are considered, edge and corner prisms: four edge prisms, each covering a subarray of four cameras [highlighted in green in Fig. 1(b)], have wedge angles  $\theta$  oriented at  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ , and  $270^\circ$  and deviate their individual fields of regard toward the edges of the global FOV. Four corner prisms, each covering two cameras [highlighted red in Fig. 1(b)], have wedge angles of  $\tan^{-1}(\sqrt{2} \tan \theta)$  (i.e.,  $\theta$  in two orthogonal directions) oriented at  $45^\circ$ ,  $135^\circ$ ,  $225^\circ$ , and  $315^\circ$  and deviate their individual fields of regard toward the corners of the global FOV. Thus there are only two prism wedge angles covering 24 cameras in total, and the central camera does not have a prism but a plate of uniform 2 mm thickness to facilitate assembly.

We use 25 nominally identical cameras (integrated package module VW6754 from ST Microelectronics) with the following basic specifications: focal length  $f = 2.96$  mm,  $f$ -number is 2.8, individual FOV is  $51^\circ \times 39^\circ$ , and sensors have  $1616 \times 1208$  pixels on a  $1.75 \mu\text{m}$  pitch Bayer matrix. Our aim is to double the FOV while also using foveated optical distortion and superresolution to yield a highly foveated image with super-resolved axial sampling and a combined global FOV of approximately  $100^\circ \times 80^\circ$ . In view of Eq. (1), we selected N-SF57 glass for the prisms, with  $n = 1.86$  and  $\theta = 16^\circ$ . The relation in Eq. (1) is plotted in Fig. 2, where it can be observed that for  $\theta = 16^\circ$  the global FOV is increased to  $100^\circ \times 80^\circ$ , as indicated by the intersection with the black lines at  $\phi_I = 50^\circ$  and  $\phi_I = 40^\circ$  for the horizontal and vertical directions, respectively (these limits are set by the sensor dimensions).

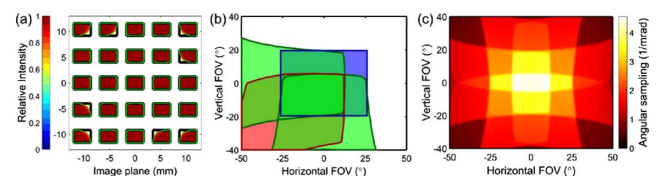
A ray-traced simulation of the system, based on the simplified assumption that the cameras are paraxial imagers, was used to estimate the angular-spatial mapping introduced by the prisms (the optical prescription of the camera lenses is not available, preventing a more rigorous ray trace). Results are shown in Fig. 3. Figure 3(a) shows the spatial variation in normalized optical intensity calculated at the image plane of all 25 sensors due to vignetting and the obliquity of the lens apertures. This small variation in relative illumination is compensated in software



**Fig. 2.** FOV enhancement from  $51^\circ \times 39^\circ$  to  $100^\circ \times 80^\circ$ . The black lines indicate the selected prism angle  $\theta = 16^\circ$  that achieves  $\phi_I = 50^\circ$  and  $\phi_I = 40^\circ$  for the horizontal and vertical directions. Right-graphs illustrate the modification of the FOV.

by flat fielding. The green boundary lines in Fig. 3(a) indicate the regions of images that contribute to the global FOV; note that some pixels, especially for corner cameras (where some regions are beyond the critical angle for which total internal reflection occurs and are not exposed), fall outside the output FOV and will therefore be discarded without implication. Conversely, Fig. 3(b) depicts the ranges of angles in object space covered by the cameras as the envelopes of the traced rays that reach the detectors (only for the central camera in blue, two edge cameras in green, and one corner camera in red to avoid excessive clutter in the figure), and Fig. 3(c) displays the resulting sampling rate of the system across the global FOV, showing that the system completely covers the desired  $100^\circ \times 80^\circ$  specification for global FOV, and also that there is a significant overlap at the central FOV, where superresolution increases the sampling frequency. The two mechanisms that contribute to increasing the foveal ratio—optical distortion introduced by the prisms and superresolution of multiple images in overlap regions—are quantified in Fig. 3(c). The angular sampling frequency is generated from a reverse ray trace from detector pixel centers to yield the density of pixels in the angular object space; the inverse of pixel density is plotted and represents the upper limit of the effective sampling rate of the system (subsampling redundancy will lower it to some extent). Superresolution is high at the center, where all 25 cameras overlap, and lower in outer regions, and, of course, enhanced resolution in the foveal region requires that optical aberrations and prism dispersion is sufficiently small that spatial frequencies above the Nyquist frequency are recorded.

A photograph of the camera array with the prism array fitted is shown in Fig. 1(d). Narrowband illumination of the scene is provided by a xenon arc lamp spectrally filtered by an interference filter with a center wavelength of 530 nm and a 10 nm FWHM. The angular dispersion introduced by the prisms at this bandwidth is equivalent to approximately 0.6 pixels, which will slightly reduce the system capacity to superresolve the

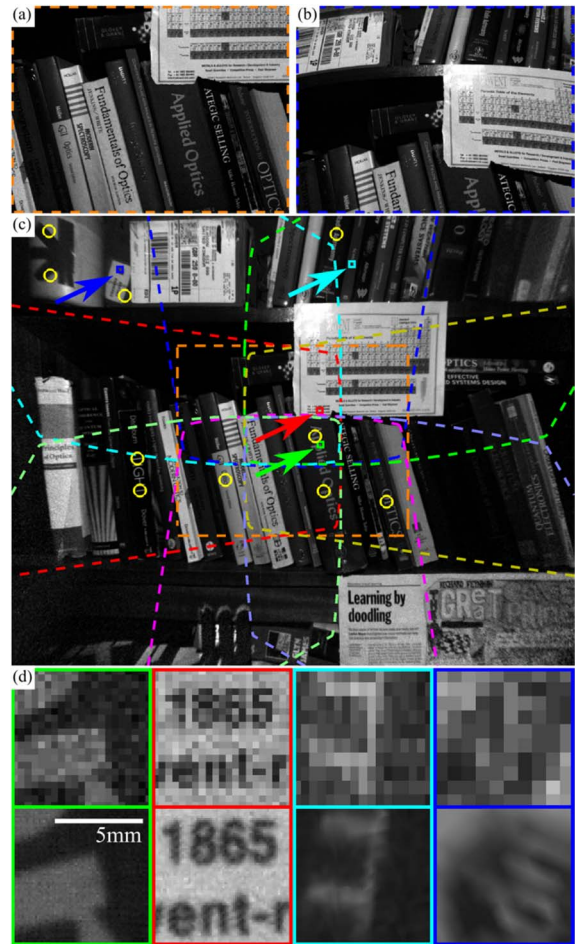


**Fig. 3.** Imaging parameters obtained from ray tracing: (a) intensity shading at the sensor plane obtained by tracing rays that sample the angular object space (relative intensity is color coded, and green lines enclose the pixels that contribute to the global FOV), (b) FOV covered by selected cameras, and (c) maximum sampling rate of the system.

scene. For broadband imaging, achromatization of the prisms is necessary, by use of compound prisms, for example [8,9]. In this proof of concept demonstration, we employed narrowband illumination to ensure that the chromatic dispersion introduced by the simple single-element prisms is sufficiently small that the point spread functions are sufficiently compact to enable superresolution to be effective. Prism-induced primary aberrations are not significant, as the cameras are designed to operate at infinite conjugate and the prisms transmit approximately collimated beams.

The image reconstruction process is composed of four steps: image projection, global image preregistration, local image registration, and interpolation. Following flat-field calibration, the recorded images are first projected onto the angular object space using a model obtained from the ray-trace simulation. This accounts for image distortion introduced by the prisms, and facilitates an approximate image registration in angular object space. Global registration and approximate geometric transformation are then performed: images from edge and corner cameras are registered with the reference center image using interest points detected and matched between image pairs using the scale-invariant feature transform (SIFT) detector [10]. However, a simple affine or projective transformation is insufficient due to imperfect agreement between actual and modeled image distortions and also due to small variations between the rotation, magnifications, and distortions of the 25 cameras. For this reason local registration is performed for small local regions (in this case we used 20x20 pixels) with respect to the reference. This local registration employs a slightly different strategy: interest points identified by the SIFT detector are cross correlated to points in the reference image, before fitting a local affine transformation. This second registration is computed with sub-pixel accuracy. The next step is to collect all registered data points (for each recorded pixel falling within the global FOV) as a scattered data cloud, which has a varying in-image density, as summarized in Fig. 3(c). Finally, we reconstruct the foveated output image by interpolating a high-resolution image from the scattered data.

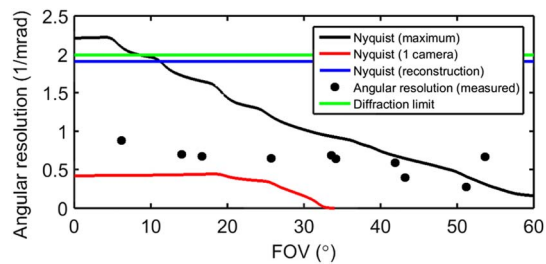
A representative reconstruction of an image of a bookcase is shown in Fig. 4. In Fig. 4(a) the individual FOV of the single central camera is shown; in Fig. 4(b) is the distorted image obtained from an edge camera (the upper-central image); in Fig. 4(c) is the global FOV reconstructed from all 25 cameras. Dashed lines in Fig. 4(c) show the region covered by each camera, and make evident how the global FOV has been doubled from the individual FOV of the central camera. Close-ups of the regions denoted by the matching colored rectangles highlighted by arrows in Fig. 4(c) are shown in Fig. 4(d) as raw images (upper row) and reconstructed images (lower row). The first two columns in Fig. 4(d) correspond to foveal areas, whereas the third and fourth columns correspond to peripheral areas covered by edge and corner cameras, respectively. Foveation effects due to both optical distortion and varying optical acuity can be observed in Fig. 4(d): although the reconstructed image has a uniform digital reconstruction sampling, the foveated variation in acuity is apparent as increased blurring in the peripheral regions. In other words, the reconstruction process achieves superresolution in the fovea and resembles interpolation toward the periphery. Comparison of Fig. 4(c) with Figs. 4(a) and 4(b) shows the extension in FOV, whereas the comparisons in Fig. 4(d) highlight the resolution increase, which is greatest in the central region of the FOV.



**Fig. 4.** (a), (b) Recorded images by central and edge cameras. (c) Reconstructed image (dashed lines show the coverage of each camera). (d) Close-ups of the arrow-highlighted regions in (c) as recorded images in upper row (cropped to approximately match the size) and reconstructed images in lower row.

We now discuss the quantification of the enhanced sampling of the image. Assuming diffraction-limited performance and ideal subpixel sampling offsets, the central FOV would be sampled at an ideal maximum angular resolution of  $\delta\varphi = \sqrt{4\phi_{ox}\phi_{oy}/(Np)}$ , where  $N = 25$  is the number of cameras and  $p = 808 \times 604$  is the number of pixels per camera (in a single color plane); the resulting extended-FOV image's resolution is then  $4\phi_{Ix}\phi_{Ly}/\delta\varphi^2 = 49$  Mpx for each color plane. This resolution is, of course, limited by subsampling redundancy, optical aberrations, registration accuracy, and ultimately by diffraction. Taking this into account and in order not to unnecessarily increase the computational and memory load by oversampling, we reconstruct results with a slightly lower sampling rate of  $4.2 \times$  the sampling rate of an individual image yielding  $\delta\varphi = 0.015^\circ$ , yielding a  $6667 \times 5333$  pixel output image.

The overall foveated imaging performance is summarized in Fig. 5. The increased sampling rate for image reconstruction can be appreciated from comparison of the black and red lines, which show the Nyquist sampling of the system and that of a single camera. The former was calculated by computing a polar average of the ideally sampled data from Fig. 3(c) (values are



**Fig. 5.** Angular resolution as recorded [from Fig. 3(c)] and reconstructed Nyquist rates. The dots are resolution measures performed at edges highlighted by yellow circles in Fig. 4(c).

calculated in terms of local Nyquist sampling for convenience in comparing angular resolution and aliasing, and are therefore halved). Note that since sensors are rectangular, the polar average produces an irregular reduction, which is why the red line starts decreasing at the vertical semi-FOV of a single camera at  $19.5^\circ$ . Empirical measures of the angular resolution performed at a range of field image points, indicated by the yellow circles in Fig. 4(c), are plotted as black dots in Fig. 5. The angular resolution here is measured by extracting the edge response function of the natural edges in the scene, calculating the line spread function by differentiation, and assigning the resolution to the angle subtended by its FWHM. (We note that for the case of a diffraction-limited system this measure matches with 84% of the Rayleigh criterion.) This angular resolution is limited by optical aberrations and diffraction for the individual cameras, and does not capture pixilation: actual angular resolution will be the lesser of optical and Nyquist angular resolutions. We note also that reconstruction artifacts will impact this resolution measure, yielding an underestimated value of the actual optical angular resolution of the system. The diffraction limit is indicated by the green line and is close to the reconstruction Nyquist sampling. Figures 3(c) and 5 assume ideal sampling offsets between the different cameras, while the system relies on randomly distributed intracamera offsets that will provide a slightly reduced effective sampling rate. Therefore, the fact that the measures of the empirical angular resolution fall between the Nyquist rate for a single camera (where aliasing is present) and the effective Nyquist rate of the system is an indication that these design parameters are in harmony, and they are indeed in good agreement with the results seen in Fig. 4. Finally, the fact that the sampling rate shown in Fig. 3(c) is not rotationally symmetric means that the foveal ratio varies with angle, and we observe a foveal ratio of 2.6 between the center field and at  $37^\circ$  mid-periphery [as a polar average, extracted from the “Nyquist (maximum)” curve in Fig. 5], and a higher ratio of 5.9 at  $50^\circ$  [in this case from a single point at diagonal, where the ratio is higher, extracted from Fig. 3(c)]. This foveal ratio arises from the varying effective sampling rates achieved and therefore will be realized only if the central resolution is indeed enabled by the optics of the cameras.

We have described this illustrative foveated imaging system in terms of angular object space assuming infinite-conjugate imaging for convenience and clarity, but for finite-conjugate imaging, it is straightforward to project angular distances to a tangent space in the required object plane, as was done for

the reconstruction in Fig. 4. The depth of field of the cameras spans from roughly 50 cm to infinity, and so the system can potentially reconstruct 3D scenes; for this, the reconstruction would need to handle registration sensitive to parallax.

An interesting felicitous benefit is obtained from the effect of the large distortions: without distortions, the nulls in the sinc form of the pixel transfer function occur at the same object-domain spatial frequency for all cameras, but the variable magnifications that occur for foveated multi-aperture imaging mean that nulls for each camera pixel transfer function occur at different object-domain spatial frequencies and increase the information and image quality recovered through superresolution [11].

In this Letter we have presented a new concept that enables multi-aperture imaging to simultaneously increase the FOV and central acuity without increasing optical complexity. The concept is based on combining independent cameras with refractive prisms to modify the individual FOV of the cameras in such a way that a higher global FOV is achieved. Increased pixel resolution is obtained in the central region of the FOV by a combination of the distortion introduced by the prism and that from superresolution of overlapping images. The optimal number of cameras and their distribution on covering the extended FOV and the optimal angles of the prisms will depend on the levels of aliasing present in the individual cameras and on required specifications, so the flexibility brought by this approach is very attractive. Overall, the concept is capable of producing low-complexity, high-resolution imaging with a varying effective resolution across a global FOV, with a significantly higher foveal ratio that can be achieved by conventional means. For broadband performance, the use of achromatized prisms adds some modest optical complexity and is tolerant to misalignment, compact, and less complex than the use of additional lens elements, as would be required for a conventional optical design. We have experimentally demonstrated the concept with a system that provides a doubling of the FOV and an increased foveal ratio of up to a factor of 6 compared to a single camera. This is thus a computational imaging solution to providing pronounced foveated imaging, of interest for applications requiring simultaneous situational awareness and vision acuity.

**Funding.** Technology Strategy Board (TSB) (KTP-8227); Engineering and Physical Sciences Research Council (EPSRC) (EP/G037523/1); Qioptiq Ltd.; ST Microelectronics Ltd.

## REFERENCES

1. G. Belay, H. Ottevaere, Y. Meuret, M. Vervaeke, J. Van Erps, and H. Thienpont, *Appl. Opt.* **52**, 6081 (2013).
2. Y. Qin, H. Hua, and M. Nguyen, *Opt. Lett.* **38**, 2191 (2013).
3. H. Hua and S. Liu, *Appl. Opt.* **47**, 317 (2008).
4. A. Ude, C. Gaskett, and G. Cheng, *Proceedings of IEEE International Conference on Robotics and Automation* (IEEE, 2006), p. 3457.
5. T. R. Hillman, T. Gutzler, S. A. Alexandrov, and D. D. Sampson, *Opt. Express* **17**, 7873 (2009).
6. B. Potsaid, Y. Bellouard, and J. T. Wen, *Opt. Express* **13**, 6504 (2005).
7. G. Carles, J. Downing, and A. R. Harvey, *Opt. Lett.* **39**, 1889 (2014).
8. N. Hagen and T. S. Tkaczyk, *Appl. Opt.* **50**, 4998 (2011).
9. G. Wong, R. Pilkington, and A. R. Harvey, *Opt. Lett.* **36**, 1332 (2011).
10. D. G. Lowe, *Int. J. Comput. Vis.* **60**, 91 (2004).
11. G. Carles, G. Muyo, N. Bustin, A. Wood, and A. R. Harvey, *J. Opt. Soc. Am. A* **32**, 411 (2015).