Superimposed multi-resolution imaging

GUILLEM CARLES,[1] JAMES BABINGTON,[2] ANDREW WOOD,[2] JASON F. RALPH,[3] AND ANDREW R. HARVEY[1,*]

[1]*School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK*
[2]*Qioptiq Ltd, St Asaph, LL17 0LL, UK*
[3]*Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, UK*
[*]*Andy.Harvey@glasgow.ac.uk*

Abstract: We present a novel approach to foveated imaging based on dual-aperture optics that superimpose two images on a single sensor, thus attaining a pronounced foveal function with reduced optical complexity. Each image captures the scene at a different magnification and therefore the system simultaneously captures a wide field of view and a high acuity at a central region. This approach enables arbitrary magnification ratios using a relatively simple system, which would be impossible using conventional optical design, and is of importance in applications where the cost per pixel is high. The acquired superimposed image can be processed to perform enhanced object tracking and recognition over a wider field of view and at an increased angular resolution for a given limited pixel count. Alternatively, image reconstruction can be used to separate the image components enabling the reconstruction of a foveated image for display. We demonstrate these concepts through ray-tracing simulation of practical optical systems with computational recovery.

References and links

1. G. Sandini, P. Questa, D. Scheffer, B. Diericks, and A. Mannucci, "A retina-like CMOS sensor and its applications," in *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop* (IEEE, 2000), pp. 514–519.
2. Y. Suematu and H. Yamada, "A wide angle vision sensor with fovea-design of distortion lens and the simulated images," in *Proceedings of IEEE International Conference on Industrial Electronics, Control, and Instrumentation* (IEEE, 1993), vol. 3, pp. 1770–1773.
3. K. Kuniyoshi, N. Kita, K. Sugimoto, S. Nakamura, and T. Suehiro, "A foveated wide angle lens for active vision," in *Proceedings of IEEE International Conference on Robotics and Automation* (IEEE, 1995), vol. 3, pp. 2982–2988.
4. X. Dallaire and S. Thibault, "Wide-angle lens miniaturization through foveated imaging," Proc. SPIE **9626**, 96261A (2015).
5. H. Hua and S. Liu, "Dual-sensor foveated imaging system," Appl. Opt. **47**(3), 317–327 (2008).
6. S. Thiele, K. Arzenbacher, T. Gissibl, H. Giessen, and A. M. Herkommer, "3d-printed eagle eye: Compound microlens system for foveated imaging," Sci. Adv. **3**(2), e1602655 (2017).
7. A. Ude, C. Gaskett, and G. Cheng, "Foveated vision systems with two cameras per eye," in *Proceedings of IEEE International Conference on Robotics and Automation* (IEEE, 2006), pp. 3457–3462.
8. Y. Qin, H. Hua, and M. Nguyen, "Multiresolution foveated laparoscope with high resolvability," Opt. Lett. **38**(13), 2191–2193 (2013).
9. G. Carles, S. Chen, N. Bustin, J. Downing, D. McCall, A. Wood, and A. R. Harvey, "Multi-aperture foveated imaging," Opt. Lett. **41**(8), 1869–1872 (2016).
10. G. Y. Belay, H. Ottevaere, Y. Meuret, M. Vervaeke, J. V. Erps, and H. Thienpont, "Demonstration of a multichannel, multiresolution imaging system," Appl. Opt. **52**(24), 6081–6089 (2013).
11. M. Mehta, E. J. Griffith, and J. F. Ralph, "Geometric separation of superimposed images with varying fields-of-view," in *International Conference on Information Fusion* (IEEE, 2014), pp. 1–6.
12. G. Curatu and J. E. Harvey, "Analysis and design of wide-angle foveated optical systems based on transmissive liquid crystal spatial light modulators," Optical Engineering **48**(4), 043001 (2009).
13. T. Martinez, D. V. Wick, and S. R. Restaino, "Foveated, wide field-of-view imaging system using a liquid crystal spatial light modulator," Opt. Express **8**(10), 555–560 (2001).
14. B. Potsaid, Y. Bellouard, and J. T. Wen, "Adaptive scanning optical microscope (ASOM): A multidisciplinary optical microscope design for large field of view and high resolution imaging," Opt. Express **13**(17), 6504–6518 (2005).
15. D. L. Dickensheets, S. Kreitinger, G. Peterson, M. Heger, and M. Rajadhyaksha, "Wide-field imaging combined with confocal microscopy using a miniature f/5 camera integrated within a high na objective lens," Opt. Lett. **42**(7),

1241–1244 (2017).

16. S. A. Burns, R. Tumbar, A. E. Elsner, D. Ferguson, and D. X. Hammer, "Large-field-of-view, modular, stabilized, adaptive-optics-based scanning laser ophthalmoscope," J. Opt. Soc. Am. A **24**(5), 1313–1326 (2007).

17. R. D. Ferguson, Z. Zhong, D. X. Hammer, M. Mujat, A. H. Patel, C. Deng, W. Zou, and S. A. Burns, "Adaptive optics scanning laser ophthalmoscope with integrated wide-field retinal imaging and tracking," J. Opt. Soc. Am. A **27**(11), A265–A277 (2010).

18. G. Carles, G. Muyo, N. Bustin, A. Wood, and A. R. Harvey, "Compact multi-aperture imaging with high angular resolution," J. Opt. Soc. Am. A **32**(3), 411–419 (2015).

19. D. B. Phillips, M.-J. Sun, J. M. Taylor, M. P. Edgar, S. M. Barnett, G. M. Gibson, and M. J. Padgett, "Adaptive foveated single-pixel imaging with dynamic supersampling," Science Advances **3**(4), e160178 (2017).

20. M. M. Mehta, E. J. Griffith, S. Maskell, and J. F. Ralph, "Geometric separation of superimposed images," in *International Conference on Information Fusion,* (IEEE, 2016), pp. 1244–1251.

1. Introduction

In traditional approaches to imaging, the choice of the focal length of the optics involves a compromise between angular resolution and field of view (FOV). In a conventional detector, the pixel detector elements are evenly spread across the FOV, and are limited in number and in extent. In many applications however it is desirable to have both high angular resolution for recognition tasks and an extended FOV to provide context and situational awareness. Simply increasing pixel count is often impractical, either due to the difficulties associated with fabricating very large pixel arrays, or due to the difficulties with processing or transmitting such high-bandwidth signals. The human visual system solves this problem by having a small central field of view with high angular resolution (the fovea) and a larger peripheral field of view with a much lower resolution. This so-called foveated approach to imaging is found in biological systems and a number of cameras based on these principles have been studied previously. In this paper, we consider an alternative model for foveal imaging, where two fields of view are superimposed on the detector array and the two images are separated by post-processing the combined image. In particular, we consider a specific optical design for such a camera, and an efficient method for extracting the two component images from the superimposed output from this camera.

Remote navigation, robot vision, remote surgery, or object tracking and recognition are all areas where a wide FOV is required for tracking and situational awareness, but a high angular resolution is also required for higher visual acuity. For example, for airborne vehicles employed in visual search and recognition roles, the angular resolution of the sensor will dictate the range at which objects can be detected and recognised, and will therefore determine the operating altitude of the aircraft. A narrow field of view will allow higher operating altitudes but will then compromise contextual awareness. One possible solution to this issue is to employ an active gimbal mechanism such that the imaging system can be scanned across the scene to build up a full situational picture. This is a conventional approach, but it comes at the cost of added weight and complexity. The ability to simultaneously image with an arbitrarily-optimized foveal ratio (i.e. the ratio of the angular resolutions of the central and peripheral regions of the FOV) would provide simultaneous contextual information and visual acuity in selected FOV regions, and therefore would be of great interest for such applications; enabling, for example, gimbal-free imaging systems.

Previous strategies to achieve foveated sampling include: the use of non-uniform sensors (with variable photoreceptor density, mimicking the variable sampling rate of the retina) [1]; optical distortion for foveated lens design [2–4]; computational integration of independent imagers with dissimilar resolutions [5–9] and the use of a single sensor segmented into multiple channels with dissimilar magnifications [10]. The high cost of hardware and the added complexity of non-uniform sensors, or the optical complexity of foveal optics design, usually make these solutions unattractive. The continuing increase in the power of low-cost computation means that a shift in complexity from optics to computation, such as is described here, is increasingly attractive. If the price per detector pixel is high, however, the cost of parallelizing independent imagers

might still be prohibitive, and segmenting a detector has the obvious drawback of reducing the number of pixels available per channel. The concept of creating a superposition of two image components, one from a narrow FOV and one from a wide FOV, in a single image was introduced in [11]. It was demonstrated that the two images could be separated by post-processing the combined image using the geometry of the superposition. Here, we adopt this approach and propose a design for a Superimposed Multi-Resolution Imager (SMRI), and we propose a novel computational imaging technique for the separation of two superimposed images. This design is an alternative solution for foveal sampling, which presents a generic dual-aperture imaging system that is able to superimpose a wide-FOV and a narrow-FOV images on a single detector, thereby achieving both a high acuity and a wide view. The image post-processing algorithm used to separate the two images is more computationally efficient than the geometric method used in [11] while providing comparable results in terms of image quality, and adding additional flexibility by allowing specific image features or spatial frequencies to be accentuated in the separated images.

The classical design approach for foveal optics exploits optical distortion to provide a specific magnification ratio across the image plane such that the detector array effectively samples the the scene in a foveated manner. This type of optical design is highly compromised for wide-FOV systems if high foveal ratios are required (corresponding in this case to the ratio of the magnifications at the center and the periphery of the image) leading to a dramatic increase in optical complexity [2–4]. A related approach is to employ a lower-complexity wide-FOV optical design and only correct for aberrations over a small region of the FOV that can be programed dynamically using spatial-light modulators [12, 13]; however, these solutions still require detectors with a very high pixel count. Alternatively, it is possible to employ two or more sensors with associated but independent optical systems to sample the scene simultaneously at different resolutions and fields of view, that is with different focal lengths, and display or fuse the recorded images as one composite image [5–8]. A different but related approach is to form a composite image from a mosaic of several images with dissimilarly optical distortions (e.g. using prisms) [9]. Computational integration of independent imaging systems is a simple, robust and powerful approach that enables arbitrary foveal ratios. However, it entails increased hardware complexity and increases the total number of detector pixels, which is particularly important when the detector is the dominant contributor to total system cost, as is the case for high-performance infrared imaging for example.

Multi-channel imaging with multiple resolutions or even employing multiple imaging modalities has been implemented for added functionality in specialized applications such as in microscopy [14], skin surface imaging [15], laparoscopic surgery [8], or ophthalmoscopy [16, 17]. An intermediate stage of this approach is to segment the detector into two or more channels of varying resolution [10]. This approach also enables arbitrary foveal ratios, however, if the width of the combined apertures fundamentally cannot exceed the sensor width, then the light-gathering and angular resolution will be limited (as sensor width then limits the f-number). It is possible to overcome this limitation to an extent however [18], but the obvious unbridgeable drawback of this strategy is that, if the sensor is segmented, the number of pixels per optical channel is reduced.

Foveated sampling can also be achieved using single-pixel cameras by implementing the appropriate structured detection patterns at an intermediate image plane that vary the way that the FOV is sampled to produce the desired variation in resolution [19]. This approach to foveation can therefore be used to reduce the time required to sample a large FOV by concentrating attention on specific areas of interest. The signal for a single frame would however then be acquired as a time series and might be suboptimal for tracking moving objects where the timing of the sampling would need to be taken into account when determining the motion of the objects being tracked.

In this paper, we propose the SMRI concept as an approach to foveal imaging that provides multi-resolution sampling of the scene by superimposing images of different resolutions onto
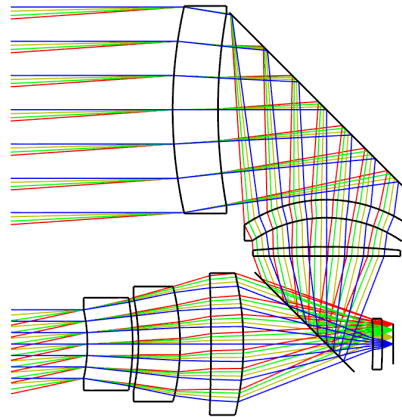
Fig. 1. Optical design of the proposed system. The design integrates a narrow-FOV (upper) and a wide-FOV (lower) channels that are superimposed on the same sensor. Each optical channel consists of three Germanium lenses. The narrow-FOV channel is folded using a mirror inserted after the first lens element, and a semireflective beam splitter combines both images onto the detector.

the same detector array. This strategy conceptually differs from previous approaches to foveated imaging. Both a wide field of view and a high-angular resolution for a smaller FOV region can be achieved simultaneously. Image separation is achieved by processing the superimposed image in a post-detection step, enabling the high-resolution image to be computationally integrated with the low-resolution wide-FOV image to yield a foveal image. This strategy relies on exploitation of the geometry of the superposition and existing redundancy in natural signals to enable digital post-detection image separation, and thus achieve a dual-FOV sampling strategy on the entire area of a single sensor. In this regard, our computational solution exploits this redundancy to increase information recorded while preserving data rates. The technique is therefore of particular significance for imaging modalities with higher associated costs per pixel. To illustrate the SMRI concept, we report on an example of a complete system and demonstrate its performance by simulation through rigorous ray tracing. The system captures the scene with a dual-FOV multi-resolution approach that provides a 3× increase in resolution in the fovea. Such specifications cannot be achieved by conventional optics without significant increases in optical complexity. The remainder of the paper is organized as follows. The optical design of the proposed system is described in Section 2. In Section 3 we model the simulation of image formation and image acquisition. Algorithms for image recovery, consisting of the separation of the wide and narrow views of the scene, are described in Section 4 and simulation results are presented. In Section 5 we discuss the benefits of SMRI to enhance simultaneous object tracking and recognition using a simulated example. In Section 6 we discuss the optimality of the foveal approach in terms of the captured information and conclude in Section 7.

2. Optical design

The layout of the proposed dual-FOV foveal system is shown in Fig. 1. The design includes a narrow-FOV channel (upper aperture) and a wide-FOV channel (lower aperture), that are combined through a mirror and beam-splitter to yield superimposed images on the sensor. The main properties and specifications are summarized in Table 1.

Optical distortion calculated by tracing the chief rays at a grid of field points is plotted in Fig. 2 for both channels. Barrel distortion for the wide-FOV channel and pincushion distortion for the narrow-FOV channel, are below 2% in both cases. The relative, field-dependent illumination for

Table 1. Specifications of the dual-FOV imaging system.

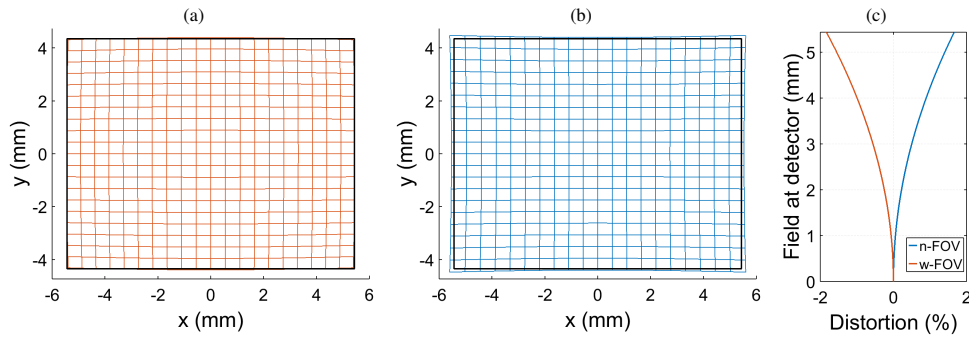| Parameter	Specification
Foveal ratio	3:1
Wide FOV (full, horizontal)	30 degrees
Wide-FOV focal length	62.2mm
Wide-FOV F/#	1.4
Narrow FOV (full, horizontal)	10 degrees
Narrow-FOV focal length	20.7mm
Narrow-FOV F/#	1.4
Spectral band	8 - 13 μm
Detector size	640×512 pixels
Pixel pitch	17 μm

Fig. 2. Geometrical distortion at the detector plane for (a) wide-FOV channel and (b) narrow-FOV channel; and (c) radial distortion plot. In (a) and (b) the thick black rectangle denotes the area of the sensor.
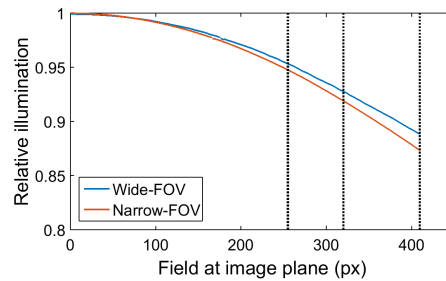

Fig. 3. Field-dependent relative illumination for both wide-FOV and narrow-FOV channels. Values are plotted as a function of radial field measured in pixels from the center of the detector array. The dotted lines denote the edge of the detector vertically, horizontally and diagonally.

both channels is plotted in Fig. 3. Both the optical distortion and the relative illumination affect together with manufacturing misalignments effect how the images are combined onto the detector, and are therefore important calibration parameters for efficient separation of the images. Finally, both channels show satisfactory optical performance as is appreciated in the modulation-transfer function (MTF) plots shown in Fig. 4.
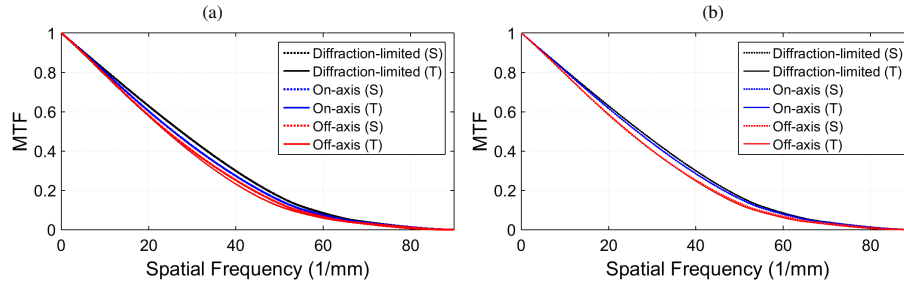
Fig. 4. Polychromatic MTFs for (a) wide-FOV channel and (b) narrow-FOV channel. In each graph the tangential (T) and sagittal (S) MTFs are ploted for diffraction-limited, on-axis and off-axis (at the edge of the detector vertically). On-axis and off-axis fields correspond with the blue and red fields traced in Fig. 1.

3. Modeling image formation and acquisition

In this section we describe simulation of the proposed system. The simulation pipeline employed for image formation and acquisition is sketched in Fig. 5. As a groundtruth for each channel, we used two video sequences with horizontal full FOV of 30 degrees and 10 degrees respectively, sampled at 5076×4056 points, corresponding to an approximate oversampling factor of 8 with respect to the pixel size of the detector, which has 640×512 pixels with a 17μm pitch. An example frame taken from a sequence is shown in Fig. 6(a) and Fig. 6(b), and simulation of imaging of a scene with a contrasted object is shown in Fig. 6(d) and Fig. 6(e).

A grid of 9×11 PSFs equally spaced within the FOV for both channels was computed to enable field-dependent convolution as simulation of image formation. PSFs were sampled at 1/8 of the pixel pitch. Each frame was transformed according to calculated optical distortion and re-sampled at 1/8 of the pixel pitch. The frame was then convolved with the grid of PSFs (bilinear interpolation was used within PSFs locations across the FOV). The convolved images were then incoherently summed accounting for the field-dependent relative illumination for each channel, and the final irradiance field was down-sampled at the detector resolution. White Gaussian noise was added to simulate detection noise with 45dB signal-to-noise ratio. The results for the selected example frames are shown in Fig. 6(c) and Fig. 6(f).

The process was repeated for all the frames in the video sequence, to generate a superimposed multi-resolution video; see Visualization 1. An interesting characteristic of SMRI is that, for a moving scene, the wide-FOV and narrow-FOV components will translate at different velocities at the detector, i.e. at the foveal ratio, as is clear from the video.

4. Image recovery

The goal of the image recovery step is to separate the wide-FOV and narrow-FOV components from each superimposed frame. One approach to this problem is to employ a recursive algorithm to solve for the intensity at each pixel location based on the known geometrical relationship of the two channels [11]. Here, we have adapted the algorithm described in [11], to incorporate optical distortion and field-dependent relative intensity; [11] dealt only with cases with no optical distortion and where the superposition coefficient in the combined image was constant. This method relates the intensity at an arbitrary location at the image plane (x, y) (centered at the optical axis) with the intensity components from the two channels as,

$$I_D(x, y) = r_W(x, y)I_W(x, y) + r_N(x, y)I_N(x, y) \qquad (1)$$

where suffices D, W and N refer to detected, wide-FOV and narrow-FOV intensities respectively, and r are the weights that account for the field-dependent relative intensity. The geometry of
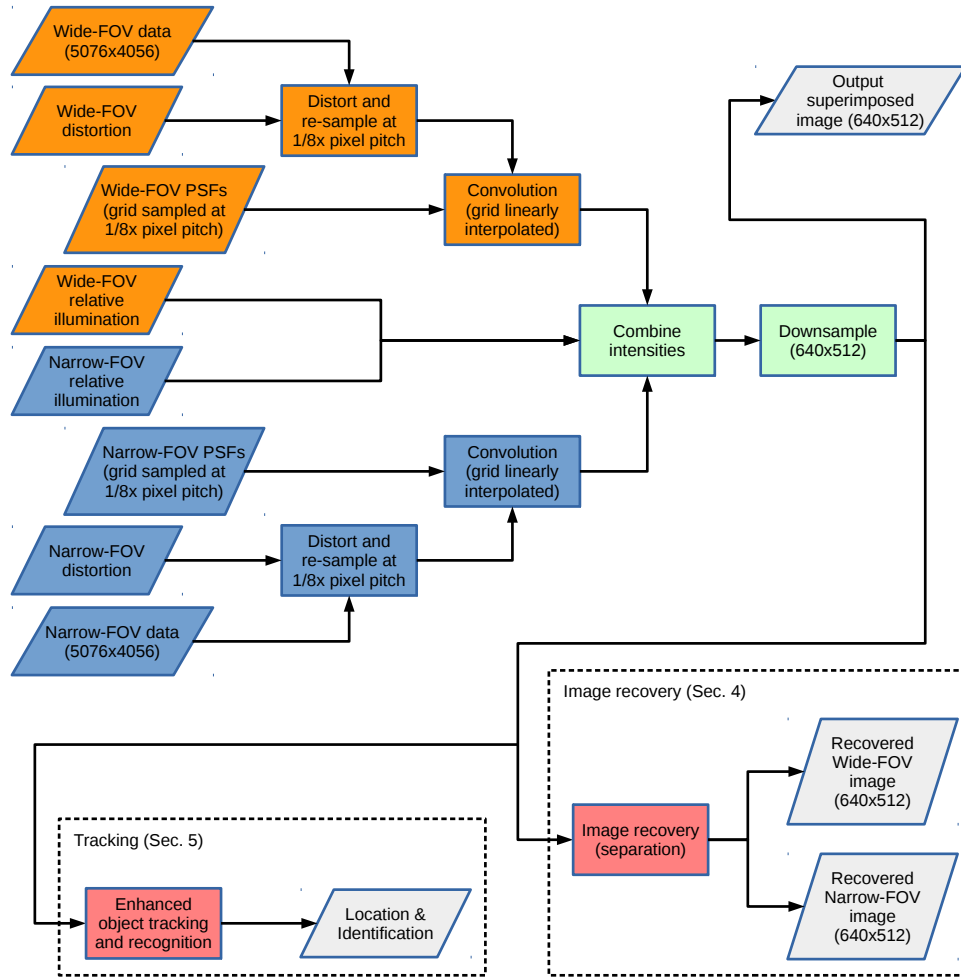
Fig. 5. Simulation pipeline of the image formation and acquisition. Postdetection image processing is also indicated with reference to Section 4 and Section 5.

the two individual images is used to develop a set of recursive equations by relating a point in the narrow-FOV scene (x_0, y_0) to the corresponding point in the (lower resolution) wide-FOV scene (x_1, y_1). The wide-FOV point is closer to the center of the image, and is combined with the intensity from a different narrow-FOV location $I_N(x_1, y_1)$, which has a corresponding wide-FOV location which is superimposed on a third narrow-FOV point $I_N(x_2, y_2)$, and so on [11],

$$
\begin{aligned}
I_W(x_1, y_1) &= I_N(x_0, y_0) \\
I_W(x_2, y_2) &= I_N(x_1, y_1) \\
I_W(x_3, y_3) &= \ldots
\end{aligned}
\tag{2}
$$

where the known information from the optical distortion enables construction of an operator $\mathcal{P}\{\cdot\}$ that performs the geometrical projection "from wide to narrow" in the image plane such that $(x_{i+1}, y_{i+1}) = \mathcal{P}\{(x_i, y_i)\}$. For convenience we explicitly employ the integer superfix (i) to refer to the location (x_i, y_i) as $I^{(i)} = I(x_i, y_i)$, and $\mathcal{P}\{\cdot\}$ ensures that $I_N^{(i)} = I_W^{(i+1)}$. Since Eq. (1) holds everywhere in the image plane, coordinates can be recursively related, asymptotically
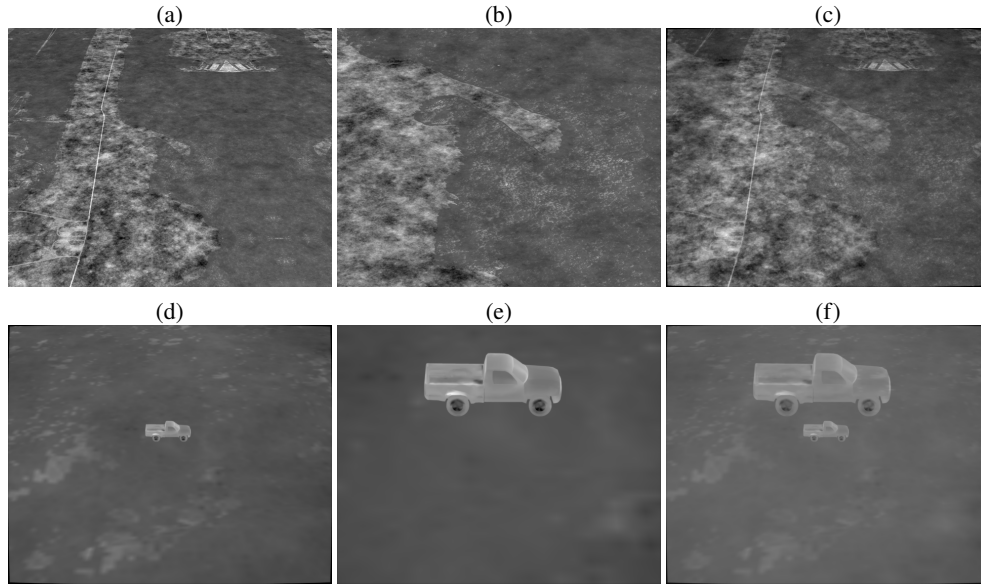
Fig. 6. Groundtruth scenes imaged by (a,d) the wide-FOV channel and (b,e) the narrow-FOV channel, and (c,f) simulation of the superimposed detection. See Visualization 1 for a video sequence for scene (a-c).

approaching the optical axis, as $(x_i, y_i) = \mathcal{P}^i\{(x_0, y_0)\}$. It then follows from Eq. (1) that

$$I_W^{(i)} = \frac{I_D^{(i)} - \frac{r_N^{(i)}}{r_W^{(i+1)}} I_W^{(i+1)}}{r_W^{(i)}} \tag{3}$$

which can be recursively solved (note that r_W, r_N are known and I_D is the measured superimposed image) assuming the initial condition $I_W^{(i')} = I_N^{(i')}$ for i' high enough that $(x_{i'-1}, y_{i'-1})$ is within one pixel area [11]. By solving Eq. (1) to find $I_W^{(0)}$ at each pixel location, the separation I_W' and I_N' can be estimated. Separation results following this method are shown in Fig. 7 referred to as the pixel-recursion method. As can be appreciated, some artifacts remain in the separated images: the recovered wide-FOV image shows some edge-artifacts caused by the narrow-FOV image that were not completely removed, and similarly the recovered narrow-FOV image has not preserved all of the high-frequency detail and therefore appears slightly blurred.

It is noteworthy that the separation of the images is an ill-posed problem, as it is intended to recover more pixels in the final image than have been recorded. This means that there are multiple solutions that will satisfy the superimposed detection condition. In other words, there will exist inherent ambiguities on whether some information content on the detected superimposed image actually belongs to the wide-FOV or the narrow-FOV views, assuming that both project equally through the system transfer function. One option to improve the separation is to use the information extracted from multiple frames and to provide a temporal average [20]. Whilst this is straightforward for a stationary camera and stationary scene, it is more complicated when the camera or objects within the scene are moving. Images must be aligned or registered before they are processed temporally. An other option is to assume a translation of the entire scene (such as would be produced if the system was implemented on a moving platform) over time and use a pair of consecutive or delayed frames to disambiguate the separation. This would be possible because the wide-FOV and narrow-FOV image components would be translated dissimilarly at the image plane.

An alternative approach uses prior information based on the statistical properties of the scene to solve the ambiguities by encouraging solutions that are statistically preferred. This approach can be implemented by building a system matrix that describes a forward model for image formation, such as the one proposed here. The motivation for this is that it provides a straightforward means to incorporate global constraints on the estimated separation. The forward model can be written as,

$$\mathbf{y} = \mathbf{D}\left(\mathbf{P}_W + \mathbf{P}_N\right)\mathbf{x} + \mathbf{e} \tag{4}$$

where $\mathbf{y}$ is the detected superimposed image I_D expressed in lexicographical order; $\mathbf{x}$ is the intensity field (in lexicographical order) that covers the wide-FOV but is sampled at the angular resolution of the narrow-FOV channel; $\mathbf{P}_W$ and $\mathbf{P}_N$ are projection matrices for the wide-FOV and narrow-FOV channels respectively that apply the respective geometrical optical distortion to the irradiance field in angular field coordinates to provide the irradiance at the image plane in spatial coordinates; $\mathbf{D}$ is a decimation matrix used to simulate irradiance integration at the pixel area; and the vector $\mathbf{e}$ accounts for noise in the detection. Eq. (4) can be inverted iteratively by using a modification of the Lucy-Richardson scheme,

$$\mathbf{x}_{n+1} = \text{diag}(\mathbf{x}_n)\left(\mathbf{DP}_W\right)^{\top}\left(\text{diag}\left(\mathbf{D}\left(\mathbf{P}_W + \mathbf{P}_N\right)\mathbf{x}_n\right)\right)^{-1}\mathbf{y} \tag{5}$$

The separation results are finally calculated by,

$$\hat{\mathbf{x}}_W = \mathbf{DP}_W\hat{\mathbf{x}} \tag{6}$$

$$\hat{\mathbf{x}}_N = \mathbf{DP}_N\hat{\mathbf{x}} \tag{7}$$

where $\hat{\mathbf{x}}$ is the estimation from Eq. (5). Results are very similar to those found from the pixel-recursion approach described above. However, it has the major advantage that it is more flexible and it enables the promotion of particular solutions with particular characteristics. As a demonstration, and to favor separations emphasizing high-frequency detail in the narrow-FOV view, we have embedded a perturbation on the iterations of Eq. (5) by adding a demagnified copy of the solution to itself using,

$$\mathbf{x}'_n = \mathbf{x}_n + \mathbf{P}_N^{\text{inv}}\mathbf{P}_W\mathbf{x}_n \tag{8}$$

where $\mathbf{P}_N^{\text{inv}}$ is the inversion of the projection matrix that maps locations at the image plane onto the corresponding field coordinates according to the narrow-FOV channel function. The effect of such a perturbation is to add a demagnified copy of the iterative solution to itself with a controlled demagnification function such that the wide-FOV components match the narrow-FOV. This perturbation has the effect of transferring information from the wide-FOV image to the narrow-FOV image, at the expense of requiring further non-perturbed iterations to recover a valid solution (we apply the perturbation once every five iterations). Results are shown in Fig. 7 (third row). The higher sharpness achieved in the narrow-FOV component is apparent, at the cost of also incorporating some content from the wide-FOV field.

Finally, we further consider the results of either the pixel-recursion algorithm or non-perturbed Lucy-Richardson but including a final sharpness-transfer process, in which high-frequency detail is copied directly to the narrow-FOV component by low-pass filtering the wide-FOV component and recalculating the narrow-FOV component from the residual from the detection, that is

$$\tilde{\mathbf{x}}_W = \mathcal{G}_\sigma\left[\hat{\mathbf{x}}_W\right] \tag{9}$$

$$\tilde{\mathbf{x}}_N = 2\mathbf{y} - \tilde{\mathbf{x}}_W \tag{10}$$

where $\mathcal{G}_\sigma$ is a Gaussian filter (σ=1.5pixels was used in this case) and results are shown in Fig. 7 (fourth row), where it can be observed as an increased sharpness of the narrow-FOV component
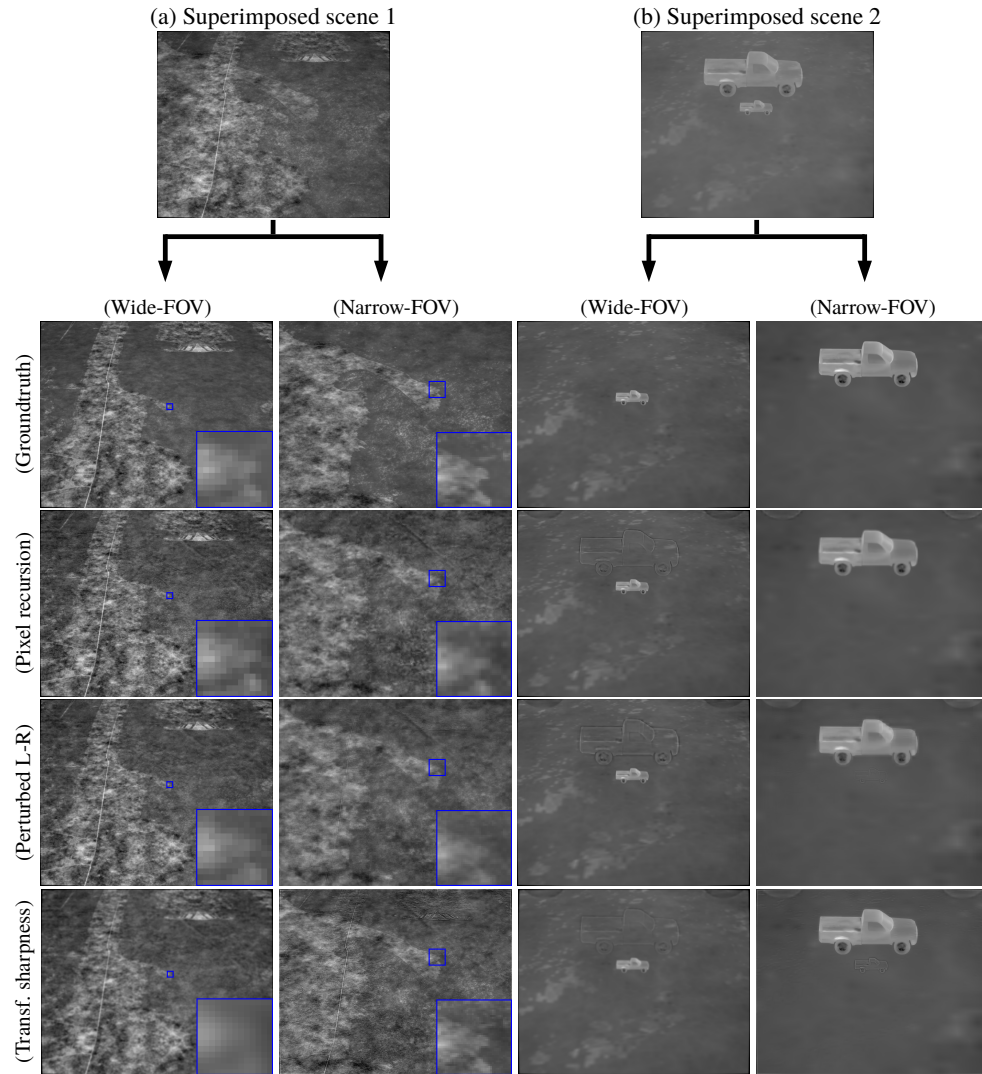
Fig. 7. Image recovery results (separation of the wide-FOV and narrow-FOV components) for the example frames shown in Fig. 6(c) and Fig. 6(d). First row corresponds to groundtruth data, second row corresponds to the pixel-recursive algorithm, third row corresponds to the system-matrix based algorithm performing the perturbed Lucy-Richardson recovery, and the fourth row results are from the sharpness transfer. The latter two are means to transfer details to the narrow-FOV reconstruction as can be appreciated in the close-up views. See Visualization 2 for a video sequence of the separation results for the scene in (a) employing the perturbed Lucy-Richardson approach.

at the expense of increased artifacts also transferred from the wide-FOV component. All in all however, it is reasonable to assume that sharpness is preferred in the narrow-FOV channel as is the component providing higher visual acuity. Note however that all are solutions compatible with the detected superimposed image when imaged through the system. In Fig. 7(a) close-up views of the region denoted by the blue square are shown for both the wide-FOV and narrow-FOV channels, highlighting the increase in angular resultion of the narrow-FOV channel.

5.   Enhanced simultaneous tracking and recognition

As is apparent from Section 4, separation of the wide-FOV and narrow-FOV components of the captured images may lead to separation artifacts that may be processed but are generally unavoidable. However, SMRI provides a means of performing object tracking and recognition with enhanced performance: the wide-FOV channel enables tracking within wider views and simultaneously the narrow-FOV channel enables object recognition or identification with increased acuity in the narrow-FOV area, and since SMRI superimposes the images onto a single sensor, there is a more efficient use of the sensor pixels overall.

In this section we demonstrate optimized and simultaneous object tracking and recognition from simulation. A video sequence that includes a static car and two moving cars was simulated. The moving cars are equal except for an identification letter imprinted in their sides. The moving cars travel across the horizontal FOV and are captured at higher angular resolution when they cross the FOV of the narrow-FOV channel. See Visualization 3 for the simulation results from the proposed system, and Fig. 8(b) for a selected frame from this sequence. A simple template-matching algorithm based on the sum of absolute differences was applied as a metric to track the cars using a model template that matches the cars at the magnification of the wide-FOV channel. See Visualization 4 to observe the metric score as the cars move, and Fig. 8(a) for the selected frame in the sequence. The detected and tracked cars are then labeled by the dashed-lined squares in red, blue and yellow. Close-ups of the tracked cars labeled in blue and red are shown in Fig. 8(c). The system is able to track the cars over the larger area covered by the wide-FOV channel, as is apparent from the peaks of the score (such as in Fig. 8(a)) in Visualization 4. In the region covered by the narrow-FOV channel, the cars are imaged at higher angular resolution. The tracked regions mapped to narrow-FOV locations are labeled by the red and blue continuous-lined squares, and are reproduced in Fig. 8(d), where it is readily seen how the identification letters are now recognizable thanks to the higher angular resolution of the narrow-FOV channel, as opposed to the close-ups in (c). In this example, both tracking and recognition were computed directly on the superimposed image.

In all, SMRI is seen to enhance simultaneous object tracking and recognition as the benefits of a wide FOV coverage and of a high-angular resolution imaging can be realized simultaneously using a single sensor, and so both functions are achieved employing fewer pixels.

6.   Discussion

The proposed concept attempts to recover two images from a single data frame. Despite the known correlation between these two images, the recovery attempts to reconstruct more pixels than are recorded. This is only possible with the assumption that recorded images show some redundancy and are sparse (or compressible) in some orthonormal basis. For example, assuming that edges in natural images are very sparse (that is, the large majority of the pixels are not edges) but carry most of the information, the detected superimposed image will capture more edge-related information because the two components will statistically produce non-overlapping edges. This concept is illustrated in Fig. 9 for a synthetic image in which the non-redundancy between the wide-FOV and narrow-FOV components happens to be perfect (i.e. there is no overlap of foreground and only a uniform background overlaps the objects) and for a natural image that also shows, statistically, an increase in edge-related information. (Of course, for the
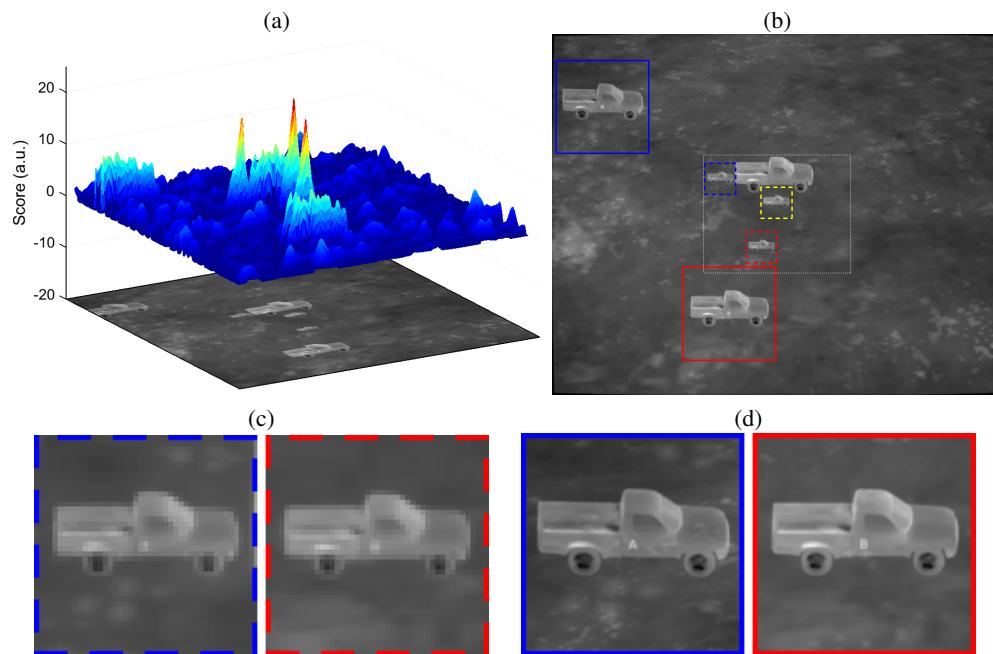
Fig. 8. Appraisal of object tracking and recognition. In the frame, three objects (cars) are within the wide-FOV view, and two move horizontally eventually within the narrow-FOV view (see Visualization 3). A template-matching algorithm based on the sum of absolute differences is used to track the cars. The metric score is plotted over the image in (a) where the peaks identify car locations (see Visualization 4 for a view of the metric score over the sequence). Peaks in metric score are used to track the cars and are labeled by the dashed-lined squares in (b) with blue, red and yellow. Close-up views of the red and blue labeled cars are reproduced in (c). The white dotted-lined rectangle in (b) shows the area of the narrow-FOV view. The location of the cars identified by blue and red are also plotted in the narrow-FOV view with the continuous-lined squares in (b), and are further reproduced in (d) where it is appreciated how the letters 'A' and 'B' in the side of each car can be recognized thanks to the higher angular resolution of the narrow-FOV channel, as opposed to the close-up views in (c).

case of the natural image, edges extracted from the superimposed image are not exactly the logical addition of the edges extracted from the individual images, since the superposition of image intensities is done prior to detection of edges, but the error is acceptably small given that the edge density is low). In both case of the synthetic and natural images, the computed fraction of edge pixels increases for the superimposed acquisition (pixel classification was done by Sobel-filtering and thresholding the images).

This discussion is emphatically relevant for object tracking and recognition: if we regard foreground pixels and edges as information of interest and discard background, then the superimposed multi-resolution system increases information captured because the wide-FOV and narrow-FOV channels may not statistically significantly compromise each other. In this latter case, tracking and recognition can be performed on the detected image directly and post-detection digital image separation would only be required for human visualization.

(a)			(b)		
(wide)	(narrow)	(superimposed)	(wide)	(narrow)	(superimposed)

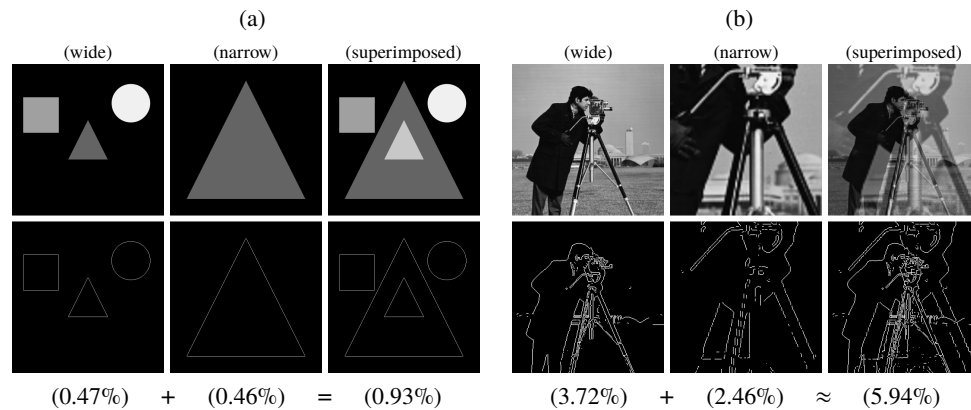(0.47%) + (0.46%) = (0.93%)   (3.72%) + (2.46%) ≈ (5.94%)

Fig. 9. Illustration of high-frequency detail content. For the synthetic image (a) there is no overlap of the edges and the information is preserved completely for both images in the superimposed image; for the natural image (b) there is some overlap but statistically the edges are not super-imposed. The underscores show the density of edge pixels (using Sobel edge detection).

7. Conclusions

We have proposed a new concept to achieve foveated imaging with arbitrary foveal ratios while making efficient use of the available pixel count. By superimposing wide-FOV and narrow-FOV images onto a single sensor, a recoverd image with both wide-view and high-acuity are achieved simultaneously. Recovery of images was simulated using superimposed images constructed using rigorous ray tracing of a realistic superimposing imaging system. In addition we also demonstrate that object tracking and recognition with optimized performance may be performed directly on the superimposed image without the need for computational separation of the wide-field and narrow-field images. These approaches exploit redundancy in the images to increase the information recorded by the detector array. Achieving image separation with reduced artifacts, either by suppressing image components that produce artifacts or by pipelined processing of sequential frames are interesting avenues for future work.

Funding