Toward Wide-Angle Microvision Sensors

Sanjeev J. Koppal, *Member, IEEE*, Ioannis Gkioulekas, *Student Member, IEEE*,
Travis Young, *Member, IEEE*, Hyunsung Park, *Student Member, IEEE*,
Kenneth B. Crozier, *Member, IEEE*, Geoffrey L. Barrows, *Member, IEEE*, and
Todd Zickler, *Member, IEEE*

Abstract—Achieving computer vision on microscale devices is a challenge. On these platforms, the power and mass constraints are severe enough for even the most common computations (matrix manipulations, convolution, etc.) to be difficult. This paper proposes and analyzes a class of miniature vision sensors that can help overcome these constraints. These sensors reduce power requirements through template-based optical convolution, and they enable a wide field-of-view within a small form through a refractive optical design. We describe the tradeoffs between the field-of-view, volume, and mass of these sensors and we provide analytic tools to navigate the design space. We demonstrate milliscale prototypes for computer vision tasks such as locating edges, tracking targets, and detecting faces. Finally, we utilize photolithographic fabrication tools to further miniaturize the optical designs and demonstrate fiducial detection onboard a small autonomous air vehicle.

Index Terms—Computational sensors, micro/nano computer vision, optical templates, optical computing, micro/nano robotics

✦

1 INTRODUCTION

THE recent availability of portable camera-equipped computers, such as smartphones, has created a surge of interest in computer vision tools that can run within limited power and mass budgets. For these platforms, the focus has been to create optimized hardware and software to analyze conventional images in a highly efficient manner. Yet, there is a class of platforms that are *still smaller*. These are microplatforms (characteristic size < 1 mm) that have power and mass constraints severe enough for large-scale matrix manipulations, convolution, and other core computations to be intractable. These platforms appear in many domains, including microrobots and other small machines [16], and nodes of far-flung sensor networks [46].

Power is the critical issue when shrinking a vision system to the microscale, with many platforms having average power budgets on the order of milli-Watts. In this paper, we present and analyze a class of microvision sensors that can help overcome the constraints of low power. Arrays of these sensors could handle a specific vision task, like face detection, as depicted in Fig. 1.

A wide field-of-view (FOV) is important for saving power because devices must either pan a low-FOV single sensor or carry multiple such sensors with different viewpoints. Our designs obtain a large FOV by exploiting the "Snell's window" effect [19], [61]. This effect, which we induce with refractive slabs, is observed by underwater divers, who see a 180 degree FOV of the outside world as grazing incident light rays are refracted at the water-air boundary by the critical angle.

Our designs also lower power consumption by reducing post-imaging computation. Template-based image filtering, an expensive component of many vision algorithms, is usually computed as a postcapture operation in hardware or software. Instead, we place attenuating templates in the optical path, allowing our sensors to perform filtering, "for free," prior to image capture. In conventional image filtering, sliding templates are applied with fixed spatial support over the image plane. Similarly, our designs ensure that the template's *angular support*, given by the solid angle ω in Fig. 1, is near constant over the hemispherical visual field. In this sense, we extend well-known planar optical filtering mechanisms [64], [41] to the wide FOV case by ensuring consistent template responses across view directions.

Our optical designs offer a new approach to efficiently implement vision algorithms on microplatforms. However, this efficiency comes at a cost, which is the penalty exacted by the mass and volume of the optics. Our main contribution is a description and formalization of the tradeoffs that exist between FOV, filtering accuracy, volume, and mass of these sensors. We discuss a variety of optical configurations, including lensless apertures, lenslets, and refracting slabs. We present solutions and tools for optimally controlling the FOV versus size tradeoff, and we validate our equations empirically.

As applications of our theory, we demonstrate a variety of sensor prototypes. We show milliscale devices, based on a web-camera platform, that are designed for edge detection, target tracking, and face detection. Results for these are demonstrated for indoor and outdoor scenes. We also demonstrate a wide-angle target tracking sensor on an embedded system with an on-board power supply. This device has an 8-bit microcontroller and shows how our
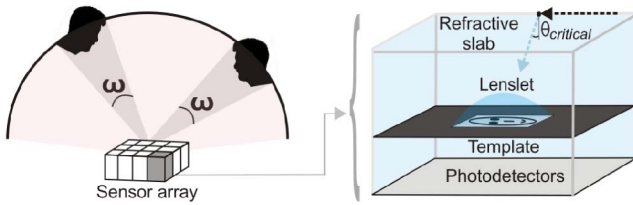
Fig. 1. We propose a miniaturized class of wide-angle sensors. Arrays of these sensors handle specific tasks. A refractive slab creates a 180 degree FOV due to Snell's law. Attenuating templates in the viewing path allow optical filtering and enable vision tasks such as locating edges, tracking targets, and detecting faces.

optical sensors can enable filtering-based algorithms on platforms with limited on-board computing power. Finally, we utilize photolithographic fabrication tools to further miniaturize the optical designs and demonstrate fiducial detection onboard a small, autonomous air vehicle.

2 RELATED WORK

Efficient hardware for micro computer vision. Our research complements work done in the embedded systems community [60], [10] because their optimized hardware and software can be coupled with our optimized optics for even greater efficiency. Indeed, all sources of efficiency should be considered to meet the power budgets available for micro platforms. For example, the successful convolution networks framework [35] was recently implemented on FPGA hardware with a peak power consumption of only 15 W [21], but this is orders of magnitude larger than what microplatforms are likely to support. Small network nodes may require an average power consumption of only 140 μW [25], [12], and microrobot peak power consumption is currently around 100 mW [29], with average power consumption around 5-10 mW [50], [59], most of it dedicated to motion.

Applied optics and computational photography. Fourier optics [24], [62] involves designing point spread functions (PSFs) of coherent-light systems to implement computations like Fourier transforms. This has limited impact for real-world vision systems that must process *incoherent* scene radiance. That said, controllable PSFs are widely used in computer vision, where attenuating templates are placed in the optical path, for deblurring [56], refocusing [42], depth sensing [36], and compressive imaging [18]. In all of these cases, the optical encoding increases the captured information and allows postcapture decoding of the measurements for full-resolution imaging or light-field reconstruction. In contrast to this encode-decode imaging pipeline, we seek optics that *distill* the incoming light to reduce postcapture processing. In this sense, our approach is closer to techniques that filter optically by either modulating the illuminating rays [44] or by filtering the viewing rays with either liquid crystal displays (LCDs) [64] or digital micro-mirror devices (DMDs) [41]. However, unlike these active, macroscale systems, we seek passive optical filtering on microplatforms.

Wide-field imaging in vision and optics. The Snell's window effect has been exploited in a classical "water camera" [61], and the projective geometry of such a pinhole camera is well understood [13]. The inverse critical-angle effect has been used to model air-encased cameras underwater [54]. In
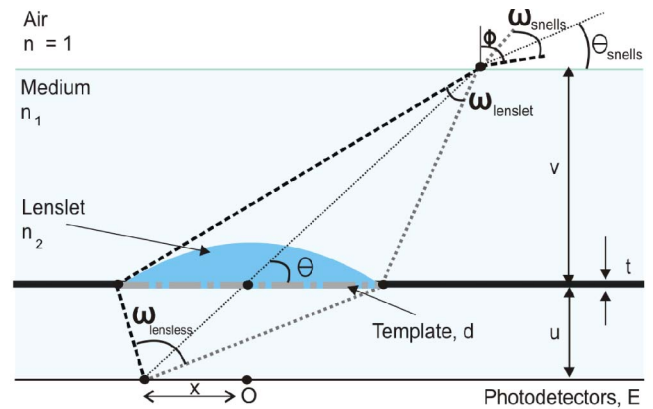

Fig. 2. *Ray diagram of our design.* By embedding a lens in a medium, we can maintain a near-constant angular support over a large portion of the frontal hemisphere.

addition to these flat refractive optical designs, a variety of wide FOV imaging systems exist in vision and optics [51], [39], and micro-optics for imaging is an active field [26], [58], [52]. While we draw on ideas from these previous efforts, our goal is quite different because we seek *image analysis* instead of *image capture*. This leads to very different designs. Our optics cannot be designed by many existing commercial ray-tracing tools (e.g., [1]) because these are created for imaging and not optical filtering.

3 DESIGN OVERVIEW AND KEY CONCEPTS

A ray diagram of our most general design, shown in Fig. 2, depicts a lenslet embedded in a refractive slab and placed over an attenuating template. All of this lies directly on top of a photodetector array, like those found in conventional digital cameras. For clarity we present a 2D figure, but since the optics are radially symmetric our arguments hold in three dimensions. (Extensions of our design to three dimensions are straightforward and Section 6 discusses one such example.) We assume that the scene is distant relative to the size of the sensor (i.e., the observed plenoptic function varies only with changes in direction and not with changes in spatial location), so the incident radiance is defined on the frontal hemisphere. We depict a single sensing element in Fig. 2, with the understanding that, for any practical application, a functioning sensor will be assembled by tiling many of these elements with complementary attenuating templates, as shown in Fig. 1. We will also assume the templates are monochromatic, but we point out that, unlike conventional postcapture image filtering, optical templates can be easily designed with task-specific spectral sensitivities. We set the embedding medium's height v to be exactly equal to the lenslet's plane of focus. While this choice seems arbitrary, it can be shown that it incurs no loss of generality when the scene is distant (see [32]).

Fig. 2 is the most general member of a *class* of designs. The refractive indices of the medium (n_1) and the lens (n_2) allow for a lensless template ($n_1 = n_2 = 1$), a template embedded in a refractive slab ($n_1 = n_2 > 1$), a template with a microlens ($n_1 = 1, n_2 > n_1$), and a template with a lens and embedding slab ($n_2 > n_1 > 1$). We will analyze all these in Section 4. Critical to this analysis is an optical filtering concept that we call the *effective field-of-view (eFOV)*. We introduce this concept next.
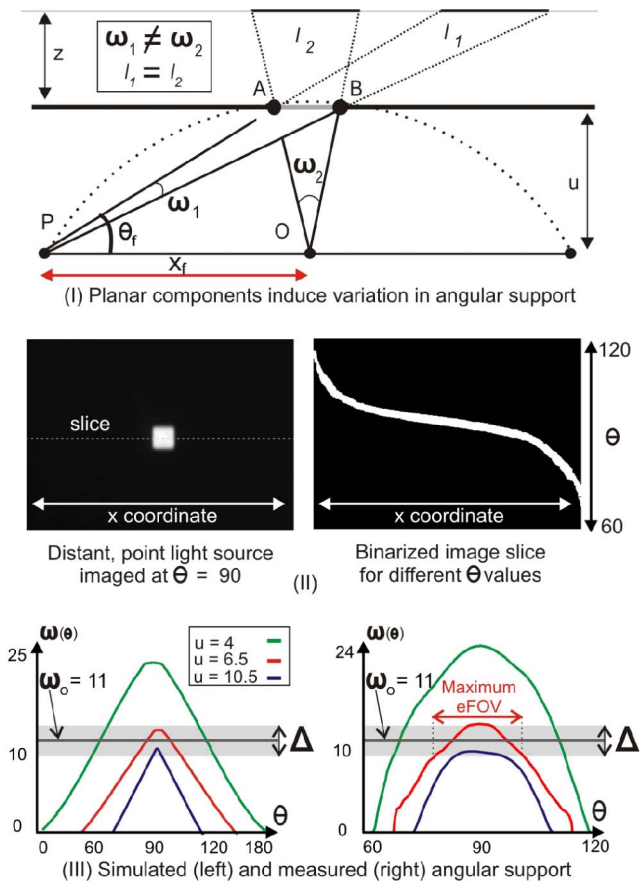
Fig. 3. *Angular support for lensless designs.* (I) shows the lensless design ($n_1 = n_2 = 1$ in Fig. 2). The angular support undergoes foreshortening with change in viewing direction and $\omega_1 \neq \omega_2$. (I) also shows the extremal photodetector location x_f whose viewing direction is θ_f. In (II), angular support is measured by observing a distant, point light source at different viewing angles θ (II left). We visualize, as a binary image created after thresholding, the illuminated pixels for a single image slice and at a particular viewing angle (II right). Integrating over the x coordinate gives the curves in (III). (III) shows measured and simulated angular support for three template heights u for a $d = 0.1$-mm pinhole. The measured angles compare well to simulations.

3.1 Effective FOV

Our design in Fig. 2 contains flat, planar components[1] that have the advantage of being readily microfabricated [9] through well-known photolithography techniques. The disadvantage of a planar construction is that it introduces perspective distortions that complicate optical filtering over a wide FOV.

To see this, consider a lensless version of Fig. 2, ($n_1 = n_2 = 1$), depicted in Fig. 3I. From similar triangles, $l_1 = l_2 = \|AB\|\frac{(z+u)}{u}$. This means that the sensor records the correlation between a scaled version of the template and a fronto-planar scene, with the effective scale of the template being determined by the distance $(z + u)$. This is the scenario for planar scenes or narrow FOV, explained in [64].

Next, consider a wide angle view of a distant scene which is hemispherical instead of planar. The system now measures correlations between the template and successive cones of light rays over the entire FOV. But because the sensor is planar, the angular support of the template, i.e.,

the solid angle that it subtends, is different for different viewing directions. For example, at point P in the figure, the angular support is ω_1. From the converse of the inscribed angle theorem, the locus of points at which AB subtends the solid angle ω_1 is a circle, shown by the dotted curve in Fig. 3I. Any other point on the photodetector array and not on the circle, such as O, has a different angular support; $\omega_1 \neq \omega_2$.

This variation in angular support is undesirable when designing a vision system. It means, for example, that a template optimized to detect targets at a particular scale will be much less effective for some viewing directions than for others. Our goal then should be to reduce such variations in angular support. To this end, we define a sensor's eFOV to be the set of viewing directions for which the corresponding angular supports are within a user-defined range. Formally, each photodetector location x in Fig. 2 defines a viewing direction θ and collects light over the angular support ω. Note that each viewing direction, θ, is contained in $(0, \pi)$. If the angular support measured in each view direction θ is represented as a scalar angular support function, $\omega(\theta)$, then we can write the eFOV as $|\Theta|$ with $\Theta = \{\theta : F(\omega(\theta), \omega_o) \leq \Delta\}$, where Δ is a user-defined tolerance and $F(\omega(\theta), \omega_o)$ is some distance metric. In the remainder of this document, we assume Θ includes the optical axis ($\theta = \frac{\pi}{2}$), and we use the L_2 distance metric so that $F(\omega, \omega_o) = \|\omega - \omega_o\|_2$.

We can *measure the eFOV* for a given physical sensor (Section 5 describes such prototypes) by sampling the angular support function $\omega(\theta)$. We do this by panning the sensor as it observes a distant point light source (Fig. 3II (*left*)). The source only illuminates pixels that collect light from a particular viewing angle. Simply counting the number of times a pixel is illuminated allows us to measure the angular support curve $\omega(\theta)$ (Fig. 3II (*right*)) and therefore the eFOV $|\Theta|$.

A concept closely related to the eFOV is the idea of *angular dot pitch*. When fabricating a template (by printing, etching, cutting, etc.) one is typically subject to constraints on the minimum realizable feature size. The distance between such features is called the minimum *dot pitch*, and this will limit our ability to shrink our optical designs. For example, if the goal is to detect faces subtending an angular support of $\omega = 2°$, and if we believe that a 20×20 template resolution is necessary to reliably detect faces of this apparent size, then the width of the template can be no smaller than 20 times the achievable dot pitch. Now, the dot pitch on the planar template will back-project to an *angular dot pitch*, which we represent by $d\omega$. In a manner similar to the variation in angular support (Fig. 3I), this angular dot pitch will vary slightly with viewing direction. However, there will necessarily be a minimal angular dot pitch value over the eFOV and this will guarantee an effective angular resolution of our optical filter. In what follows, we will assume that both the desired angular support ω_o and the angular dot pitch $d\omega$ exist as user-provided specifications.

4 ANALYSIS

With the concept of eFOV in hand, we can analyze the class of sensing elements shown in Fig. 2 and understand

---

1. Curved sensors remain in nascent development [31] and Section 6 discusses possible designs that use them.

TABLE 1
Summary of Symbols Used in the Analysis

Variable name	Meaning
Π	**Set of design parameters**
u	Template height above photodetector array
d	Template width
n_1	Refractive index of medium
n_2	Refractive index of lenslet
R	Radius of curvature
Ξ	**Angular filtering specifications**
ω_o	Desired angular support
Δ	Tolerance for angular support
F	Distance metric between angular supports
$d\omega$	Angular dot pitch
Ψ	**Set of design constraints**
d_{min}	Minimum template width
E_{max}	Maximum photodetector length
$eFOV$	Effective field of view (informal)
Θ	Effective field of view (formal)
ω_{snells}	Angular support for refractive slab
$\omega_{lensless}$	Angular support for lenslet
$\omega_{lenslet}$	Angular support for lensless design
x	Photodetector array location
ϕ	Largest incidence angle
θ	Viewing direction
θ_{snells}	Viewing direction after refraction through slab
$\omega(\theta)$	Angular support as a function of viewing angle
x_f	Extremal photodetector location
θ_f	Extremal viewing direction (at x_f)
O	Origin of design
f	Lenslet focal length
v	Height of focal plane above template
ρ_1	Density of n_1
ρ_2	Density of n_2
V	Design volume
W	Design weight
Ω	Angular support in three dimensions
t	Aperture thickness
ω_{vig}	Reduced angular support due to vignetting

the tradeoffs between the eFOV and the element's volume and mass.

A single sensor element' design parameters (Fig. 2) form a five-dimensional vector $\Pi = \{u, d, n_1, n_2, R\}$, where u is the template height, d is the template width, n_1 is the medium's refractive index, n_2 is the lenslet's refractive index, and R is its radius of curvature. Note that the eFOV depends only on angular quantities (angular support ω, viewing direction θ), which are invariant under uniform scaling of the lengths and distances in Fig. 2. This means there exists *at least* a one-dimensional family of lensless design parameters, $\Pi_k = \{ku, kd, n_1, n_2, kR\}$, parameterized by scale k, that have identical eFOV.

Given a set of user defined angular filtering specifications $\Xi = \{\omega_o, \Delta, F, d\omega\}$, selecting the design parameters Π determines both the angular support function, $\omega(\theta)$ (and therefore the eFOV), as well as the physical extent of the optics (its volume and mass). How do we go about finding the "right" design parameters Π? In the following sections, we will derive equations and present empirical analysis, in the form of a look-up table, to answer this question. Table. 1 summarizes our notation.

Design constraints. The design parameters Π are limited by a number of constraints, which we denote by Ψ. Here, we list all types of constraints Ψ for completeness. However, we only use a clearly defined subset of these during the analysis. There are two classes of constraints: 1) The design parameters Π must be *physically plausible*, with

$u, d, R \geq 0$, $n_1, n_2 \geq 1$, $d \leq 2R$ (from the lens equation), and $n_2 \geq n_1$ (convex lens); 2) the design parameters Π must allow *easy microfabrication*. The second class of fabrication constraints relate to the minimum size of physical features that can be reliably constructed. Our ability to shrink the design can be limited by the minimum template width d_{min} for which the realizable dot pitch achieves the desired angular dot pitch $d\omega$, as explained previously, the maximum photodetector array length E_{max} that can be afforded, or the minimum aperture thickness t, whose vignetting effect on angular support is explained in Section 6. These constraints will relax as fabrication processes evolve, but since our analysis is based on geometric optics, there currently exists a strong lower bound on size induced by diffraction [38].

4.1 Lensless Design in Air

Consider a lensless version of Fig. 2 with refractive indices $n_1 = n_2 = 1$, implying that the design parameter space is two-dimensional, $\Pi = \{u, d\}$. In this case, the angular support of the template is equal to $\omega_{lensless}$ in Fig. 2, and for notational convenience we represent this by ω for the remainder of this section. We define x_f as the extreme point furthest from the origin O with θ_f the corresponding extreme view direction (Fig. 3I). Since the lensless configuration has no optics, mass is negligible for this design. We can define an "optimal design" as the one that achieves the largest possible eFOV while fitting within the smallest possible volume, given by (in this 2D case) $2ux_f$.

Consider a point on the photodetector array at a distance x from the origin O, as shown in Fig. 2. We use the cosine law to obtain an expression for the angular support ω. To calculate the sides of the triangle whose vertex is at this point, we construct a perpendicular to the template, which is of length u, the template height. This gives two right triangle expressions for two hypotenuses, $u^2 + \left(\frac{d}{2} - x\right)^2$ and $u^2 + \left(\frac{d}{2} + x\right)^2$. Using these with the cosine law, and since $x = u \cot \theta$, we can obtain an expression for the angular support function:

$$\omega(\theta) = \arccos\left(\frac{2u^2 + 2(u\cot\theta)^2 - \frac{d^2}{2}}{2\sqrt{\left(u^2 + \left(\frac{d}{2} - u\cot\theta\right)^2\right)\left(u^2 + \left(\frac{d}{2} + u\cot\theta\right)^2\right)}}\right).$$

(1)

To understand how the angular support $\omega(\theta)$ changes as the design parameters $\Pi = \{u, d\}$ are varied, we directly measured angular support curves $\omega(\theta)$ (using the procedure described in Section 3.1) for a fixed template width $d = 0.1$ mm and three different template heights $u = \{4, 6.5, 10.5\}$ (Fig. 3III). These experimental curves matched the theoretically expected angular support curves from (1). Note that the angular support curves are symmetric because $\omega(\theta) = \omega(\pi - \theta)$ in (1).

A user-specified target angular support ω_o and tolerance Δ define a region that is marked as a gray bar in Fig. 3III. The central, red curve in Fig. 3III is contained inside the gray bar for the larger interval of viewing angles and, therefore, has a higher eFOV.

Given the general shape of the curves in Fig. 3III and our assumption that the optical axis $\theta = \frac{\pi}{2}$ is included in the eFOV, we can intuitively describe a design that

maximizes the eFOV by having the angular support curve $\omega(\theta)$ be tangential to the horizontal line $\omega = \omega_0 + \frac{\Delta}{2}$ at $\theta = \frac{\pi}{2}$ (similar to the red curve). Substituting these values into (1), we obtain

$$\cos\left(\omega_o + \frac{\Delta}{2}\right) = \frac{2u^2 - \frac{d^2}{2}}{2u^2 + \frac{d^2}{2}}. \tag{2}$$

Using the fact that the template width d and height u must be positive, we can rewrite (2) in the form:

$$u = \frac{d}{2}\sqrt{\frac{1 + \cos\left(\omega_o + \frac{\Delta}{2}\right)}{1 - \cos\left(\omega_o + \frac{\Delta}{2}\right)}}. \tag{3}$$

Equation (3) provides a necessary condition that must be satisfied by u and d in order for the angular support function to be tangent to the upper-bound line at $\theta = \frac{\pi}{2}$ and therefor, have the maximal eFOV. We also observe that u and d are linearly related and therefore invariant under global scaling, as expected.

The above discussion suggests a two-step algorithm for finding the optimal lensless design: 1) Arbitrarily select the template width d and compute the corresponding u from (3); 2) globally scale the design parameters Π downward such that the constraint $d \geq d_{min}$ is satisfied. We only consider physical constraints and the minimum template width d_{min} in the full set of constraints Ψ, but the same procedure can be used directly for other constraints. The optimal design after global scaling is denoted as $\Pi^* = (u^*, d^*)$.

The volume and eFOV of this optimal design Π^* can be determined analytically. To see this, note that we want every point on the photodetector to have an angular support within the gray bar in Fig. 3III (we do not want any "wasted" pixels). Therefore, the angular support of Π^* should behave as the red curve in Fig. 3III; when the curve exits the gray bar region, the corresponding viewing ray should be the extremal ray (Fig. 3I) such that $\theta = \theta_f$ and $\omega_f = \omega_o - \frac{\Delta}{2}$. By substituting into (1), and using (3):

$$C = \frac{K + K(\cot \theta_f)^2 - 1}{\sqrt{\left(K + \left(1 - \sqrt{K}\cot \theta_f\right)^2\right)\left(K + \left(1 + \sqrt{K}\cot \theta_f\right)^2\right)}}, \tag{4}$$

where we denote for convenience $K = \frac{1 + \cos(\omega_0 + \frac{\Delta}{2})}{1 - \cos(\omega_0 + \frac{\Delta}{2})}$ and $C = \cos(\omega_0 - \frac{\Delta}{2})$. The above expression can be rewritten as an easily solvable biquadratic equation in terms of $X = \cot \theta_f$, as follows:

$$\begin{aligned}
&\left(C^2 K^2 - K^2\right)X^4 \\
&+ \left(2C^2 K^2 - 2C^2 K - 2K^2 + 2K\right)X^2 \\
&+ \left(C^2 K^2 + C^2 + 2KC^2 - K^2 - 1 + 2K\right) = 0.
\end{aligned} \tag{5}$$

Ignoring complex solutions, (5) has at most two pairs of solutions $X = \pm X_i$, $i = 1, 2$. From each such pair, we obtain supplementary angles $\theta_f = \text{arccot}(X_i)$ and $\pi - \theta_f$, corresponding the left and right extreme points of the photodetector array (Fig. 3I), and therefore representing the same design. Each such solution of (5) completely characterizes the maximum eFOV as $\Theta = (\theta_f, \pi - \theta_f)$. Interestingly, the

actual value of the maximum eFOV depends only on the user defined parameters ω_o and Δ.

We will now prove that only one solution pair of (5) is physically meaningful. From the converse of the inscribed angle theorem, the locus of the points at which the template subtends an angle ω_f is a unique circle, with the template as a chord of length d. This circle may only intersect the photodetector array line at most at two points. Therefore, there can be *at most* two solutions for X. Additionally, since the angular support is symmetric, continuous (see Fig. 3III), has a maximum of $\omega_o + \frac{\Delta}{2}$ (3) and a minimum of 0 (the limit as θ approaches 0 in (1), (4) will be satisfied at least twice when the angular support becomes $\omega_o - \frac{\Delta}{2}$ and there are *at least* two solutions for X. Therefore, (5) has exactly one pair of physically consistent solutions (the inconsistent solution pair occurs when (4) is squared) which uniquely defines the maximum eFOV.

In summary, we select the optimal design $\Pi^* = (u^*, d^*)$ with the help of (3). The maximal eFOV of Π^*, which are the angles contained in $(\theta_f, \pi - \theta_f)$, is uniquely obtained from (5), which is biquadratic and has well-known, closed-form solutions. The volume of the optimal design Π^* is $2u^* x_f$, where $x_f = u^* \cot \theta_f$.

4.2 Lenslet Design in Air

For a lenslet in air, the lenslet's refractive index is higher than the surrounding medium (air), $n_2 > n_1 = 1$, and therefore the design parameters are $\Pi = \{u, d, n_2, R\}$. In this case, the angular support of the sensor is equal to $\omega_{lenslet}$ in Fig. 2, and in the discussion below the symbol ω will refer to this angle.

Plano-convex lenses do not have a favored orientation, and we can use a downward facing lenslet, as in Fig. 4, where the lens lies between the template and the photodetector array. This configuration has the advantage that the lenslet does not add any extra volume to the design, and the volume is calculated exactly as in Section 4.1. Such a lenslet adds a physical constraint of $n_2 \leq 2$, which forces the radius to be less than both the focal length and template height $R \leq f < u$.

We also note that, unlike in the lensless case, the lenslet design cannot be assumed to be massless, and we must take into account the weight of the lenslet. This is calculated by multiplying the lenslet volume, computed as a spherical cap, with the density corresponding to the refractive index n_2, obtained by assuming a linear relationship between optical and physical densities [22].

We propose a two-step algorithm to obtain the design parameters Π: 1) Find a corresponding lensless design $\Pi_l = \{u_l, d_l\}$, and 2) trade off volume and weight using lenslet parameters (n_2, R).

Before we explain the algorithm, we first demonstrate that, for any lenslet, there exists a corresponding lensless design with identical eFOV. To see this, consider the angular support equation for the lenslet, obtained in a manner similar to that of the previous section, from similar and right triangles in Fig. 2:
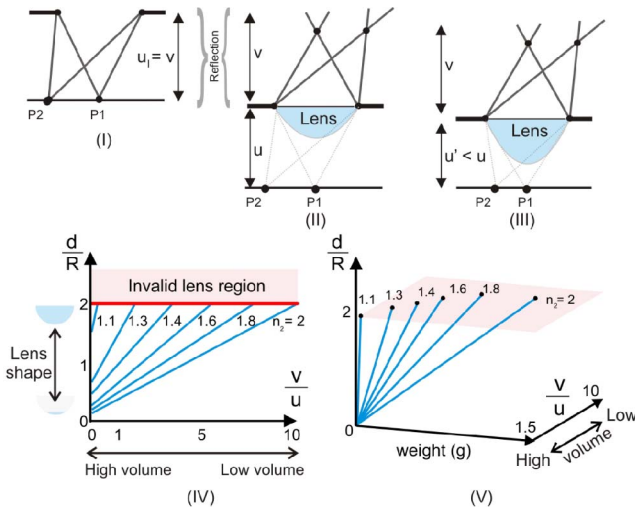
Fig. 4. *Volume-weight tradeoffs for lenslets in air*. The ray geometry in (I) is identical to the unrefracted, incident rays in (II). The design in (III) is heavier that that in (II), but requires a smaller volume ($u' < u$). Reducing the volume by increasing the refractive index (IV) has a cost in increased weight (V). Valid thin lenses must have $d \leq 2R$.

$$\omega(\theta) = \arccos\left(\frac{2v^2 + 2(v\cot\theta)^2 - \frac{d^2}{2}}{2\sqrt{\left(v^2 + \left(\frac{d}{2} - v\cot\theta\right)^2\right)\left(v^2 + \left(\frac{d}{2} + v\cot\theta\right)^2\right)}} \right).$$

(6)

By comparing (1) and (6), we observe that a lensless design with template width d and template height v would have identical angular support and, therefore, identical eFOV. Figs. 4I and 4II illustrate this idea with an intuitive geometric argument. The ray geometry in (I) is the same (under mirror reflections) to the exterior, unrefracted rays in (II). Further confirmation is provided in Fig. 5II, which shows simulated and measured angular support curves $\omega(\theta)$ for a 3-mm lenslet, which are similar in shape to those described in the previous section for lensless designs.

Returning to step 1 of our algorithm, given angular filtering specifications Ξ and constraints Ψ, we first find the best lensless design parameters $\Pi_l = \{d_l, u_l\}$ using the 2-step algorithm in Section 4.1. As discussed previously, this provides the largest possible eFOV and the lowest volume.

Next, from the argument above, we generate a lenslet design $\Pi = \{u, d, n_2, R\}$ with identical eFOV to Π_l as follows: First, we set $d = d_l$; then, we use the thin lens equation with $v = u_l$ and $f = \frac{R}{(n-1)}$ to obtain

$$u = \frac{Ru_l}{u_l(n_2 - 1) - R}.$$

(7)

In the above, we have, for the moment, arbitrarily selected values for the refractive index $n_2 > 1$ and a valid radius $R \geq \frac{d}{2}$. We note that there is no linearly scaled lenslet design $\Pi_k = \{ku, kd, n_2, kR\}$ with both higher eFOV and lower volume than Π. This is because, by design, we created Π from the best lensless design Π_l in the one-dimensional family of scaled lensless designs. However, there could be nonlinear changes to Π that could lower the weight and volume, while keeping the eFOV the same.

In step 2 of our algorithm, we perform such nonlinear manipulations to the design parameters Π. This is done by keeping the template width d and plane of focus v fixed, and changing the three remaining design parameters, u, n_2 and R. Due to the constraint of (7) these manipulations correspond to only two degrees of freedom. From the equation, we note that decreasing u (to lower the design volume) implies either reducing R (a larger, "rounder" lens) or increasing n_2 (a denser lens). Therefore, lowering the volume results in increasing the lenslet weight and the two-dimensional parameter space represents a volume-weight tradeoff. Consequently, it is impossible to obtain an eFOV-maximizing design that has both the lowest weight and volume.

Fig. 4 illustrates this tradeoff, for a desired angular support of $\omega_o = 16°$. The graphs in Fig. 4IV are the volume reductions achieved by different refractive indices. The best compression is obtained where these lines intersect the $d \leq 2R$ constraint in Ψ. However, Fig. 4V shows the
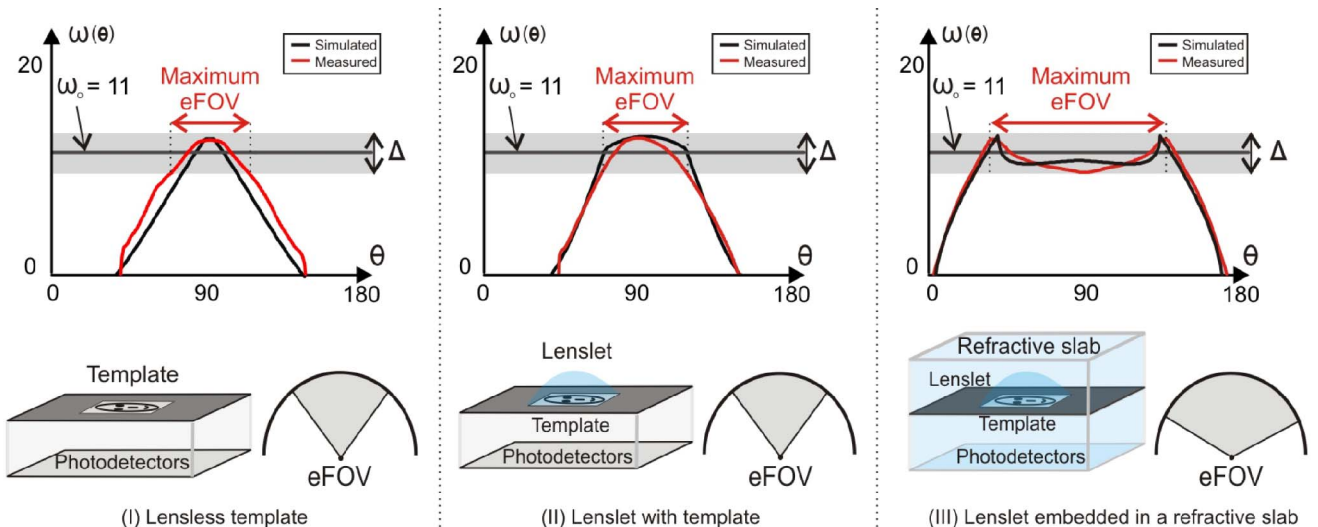

Fig. 5. *Simulated and measured angular support graphs for lensless sensors, lenslets in air, and embedded lenslet sensors*. The eFOVs are given by the set of angles θ for which $\omega(\theta) \in \omega_o \pm \frac{\Delta}{2}$. Note the high eFOV of the embedded lens with the Snell's window effect.

corresponding increases in weight as the volume decreases, suggesting that, unlike the lensless case, there is no "best" choice, but a space of designs from which one can make an appropriate choice for a given platform.

4.3 Designs with Snell's Window

For a design with snell's window, there is either a lensless template embedded in a medium $n_2 = n_1 > 1$, or a lenslet embedded in a slab $n_2 > n_1 > 1$. In the discussion below, ω will refer to ω_{snells} in Fig. 2 and the design parameters are $\Pi = \{u, d, n_1, n_2, R\}$.

Inside the medium, the relationship between the embedded lensless template and the embedded lenslet is similar to that of Sections 4.1 and 4.2. For example, for any lenslet embedded in a medium, $\Pi = \{u, d, n_1, n_2, R\}$, we can find an equivalent embedded lensless template design $\Pi_l = \{u_l, d_l, n_1\}$ using $d_l = d$ and using a version of (7), taking into account the change in effective lenslet focal length due to the embedding [45]. Since the design issues within the medium are similar to previous sections, we will only concern ourselves here with the air-refractive slab boundary.

In Fig. 2, each photodetector location x collects light rays. One of these rays has the largest incident angle, which we denote by ϕ. If we increase the distance of x from the origin O, then the corresponding largest incident angle ϕ increases. However, the maximum value of ϕ is bounded by the critical angle $\arcsin \frac{1}{n_1}$. Beyond this point, further increases in the photodetector distance x will result in the "cropping" of the template; only a portion of the template will be illuminated by incident light rays from the scene.

The angle ϕ is determined by all five parameters in the design Π. Since we wish to deal only with the air-surface boundary effects, this makes ϕ useful as a proxy for design parameters within the medium. Using only ϕ, the viewing angle θ_{snells} and the refractive indices, along with similar triangles and right triangle equations, we can derive the following expression for the angular support ω:

$$\frac{\sin(\|\omega - \phi\|)}{\sqrt{(n_1)^2 - \sin^2(\|\omega - \phi\|)}} + \frac{\sin(\phi)}{\sqrt{(n_1)^2 - \sin^2(\phi)}} - \frac{2\cos(\theta_{snells})}{\sqrt{n_1^2 - \cos(\theta_{snells})^2}} = 0. \quad (8)$$

Derivation details are available as supplementary material in [32]. Empirically, we have found that the above equation can be considered as an implicit function for $\omega(\theta_{snells})$. We numerically solve (8) by evaluating a one-dimensional search for ω (for each value of θ_{snells}), and Fig. 5III shows (in black) a curve for particular angular filtering specifications Ξ. Note that the shape of the angular support function differs greatly from those of the lensless and lenslet designs in air. In particular, it remains within the tolerance bounds, defined by the Δ, for a larger set of viewing angles than for the nonembedded cases. Additionally, the angular support curve shows a discontinuity. This occurs exactly when ϕ becomes the critical angle, and results in both the cropping of angular support and a sharp fall in the curve. We have performed experiments to verify this behavior (curve shown in red) and found that it
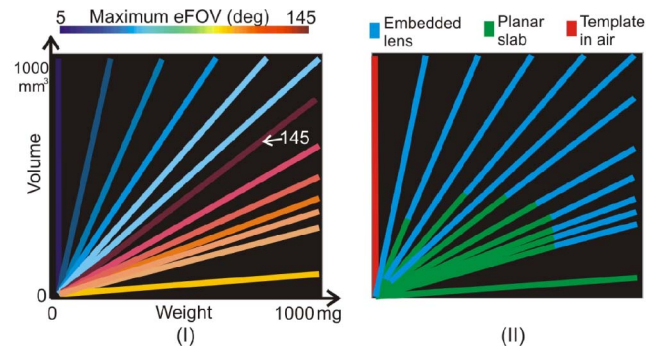

Fig. 6. *Volume-Weight lookup table for* $\omega_o = 12°$. Here, we project the (Volume, Weight, eFOV) look-up table onto the Volume-Weight plane, by only plotting the maximal eFOV at each plane coordinate. Note that design parameters Πs with the same eFOV form one-dimensional spaces (lines). However, more than one configuration can create the same eFOV, as shown by the masks on the right, which color-code the optical designs. The design variations in this figure are best viewed in color.

matches the theory. This demonstrates that using an embedding medium increases the eFOV.

The sensor's volume, determined by the design parameters Π, can be written as $V = 2x_f u$. Its weight is given by $W = V_l \rho_2 + (V - V_l)\rho_1$, where V_l is the volume of the lenslet, computed as a spherical cap, and ρ_1 and ρ_2 are the densities of the refractive media with indices n_1 and n_2. As before, we obtain these by assuming a linear relationship between optical and physical densities [22].

Like the lenslet in air, there is no "best" design, but a design space that allows trading volume and weight for a particular eFOV. Unlike the two previous cases, the Snell's window designs do not have analytic solutions for the eFOV. Applying numerical solutions to (8) for different design parameters Π suggests an empirical strategy for exploring the design space, which we discuss further in the next section.

4.4 Lookup Table for Sensor Designs

Consider now design parameters Π that encompass all previously discussed scenarios. Using the analysis of the previous sections, we provide an empirical overview of the design parameters $\Pi = \{u, d, n_2, n_1, R\}$ and build a look-up table for designers wishing to constrain or specify the desired weight, volume, and eFOV characteristics of a sensor. We take advantage of the small sensor sizes and assume reasonable ranges on the values of u, d, and R. For every set of design parameters Π within this range, we find the eFOV. For the lensless and lenslet designs in air, we can take advantage of the analytic solutions, whereas for the Snell's designs we use grid-based numerical evaluations. Formally, for a given set of angular filtering specifications Ξ, by densely sampling the physically plausible part of the parameter space Π and computing $(V, W, eFOV)$ for each sample, we produce a (one-to-many) map:

$$m_\Xi : (V, W, eFOV) \rightarrow \Pi. \quad (9)$$

This can be used by designers to choose sensor materials and physical dimensions that meet the volume and/or weight constraints of their platform while providing the desired angular filtering characteristics Ξ.
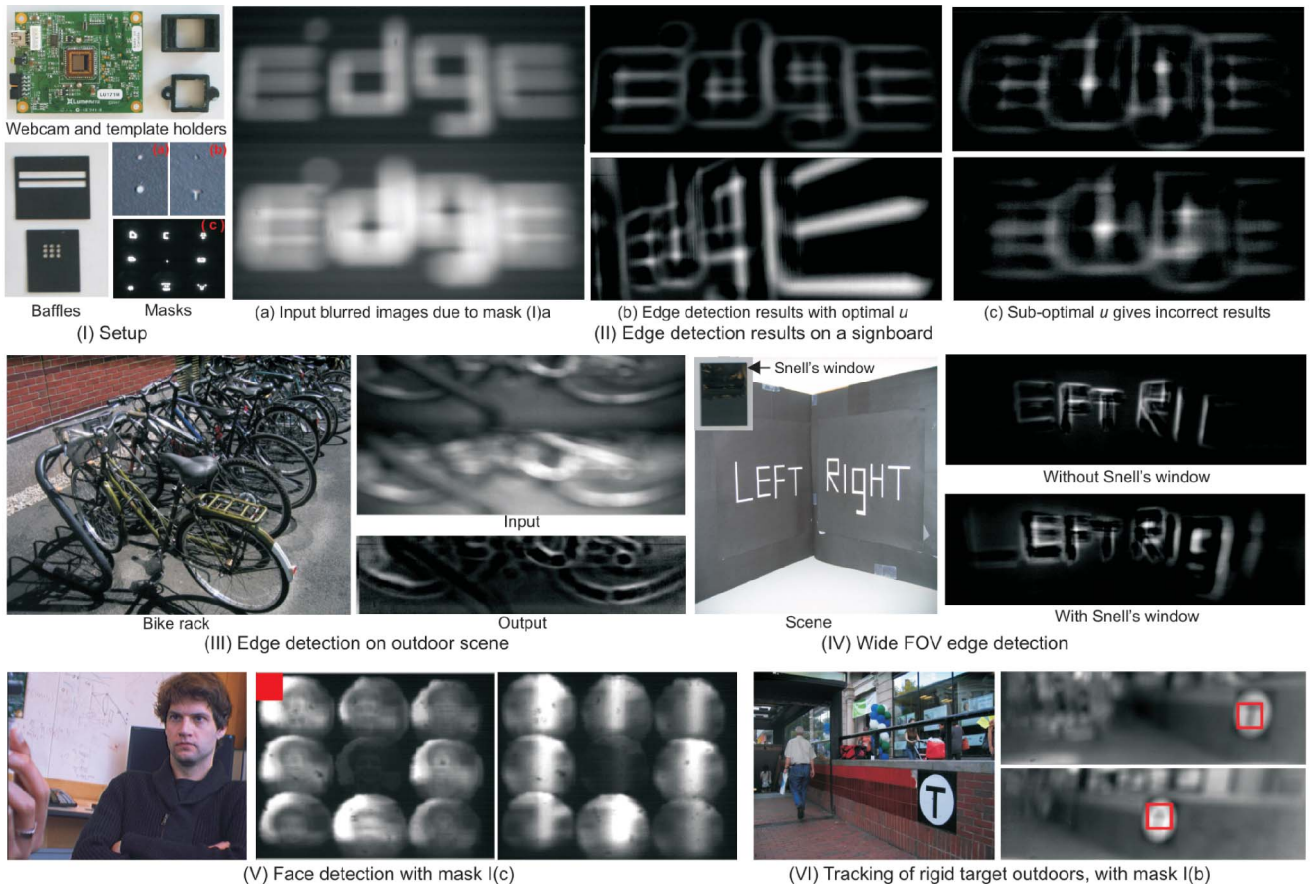
Fig. 7. *Applications*. In (I), we show our setup: a camera with custom template holders. We use template I(a) to obtain two blurred versions of the scene, as in II(a). This allows edge detection through simple subtraction, as in II and III. Without our optimal parameters, the edge detection is unreliable II(c). Wide-FOV edge detection is possible with a Snell's window enhanced template shown in (IV). In (V), mask I(c) was learned from a face database [34], and the nine mask responses are used by a linear classifier to provide face detection. In (VI), we show rigid target tracking using mask I(b), which includes two templates. More results are available at [32].

One way to visualize this map is to determine the maximum-possible eFOV for each volume-mass pair by computing $eFOV_{max}(V, W) = max_{eFOV}(V, W, eFOV)$. Fig. 6 shows such a visualization for a desired angular support of $\omega_o = 12°$ and a user-defined tolerance $\Delta = 2.4°$. Each point in the plane shows the maximal eFOV of all sampled design parameters Πs at that point. Not every set of parameters Π was sampled, and designs that were not included create black spacings. In Fig. 6I, we color code the graph according to eFOV, clearly showing lines with the same eFOV. This is because, given any set of design parameters Π, we can generate a family of designs with equivalent eFOV through $\Pi_k = \{ku, kd, n_2, n_1, kR\}$. However, unlike in previous discussions, there may exist other optical designs, outside this one-dimensional space, that have the same eFOV. Reddish hues in (I), corresponding to higher eFOV, slope toward higher weight, implying that heavier refractive optics enable larger eFOV, as expected. Each point $(V, W, eFOV_{max})$ maps to a point in the parameter space Π that can be one of the three types. This is depicted by the color transitions (lensless as red, lenslet as blue, snell's as green) in some lines in Fig. 6II. The red vertical lensless design in Fig. 6II is likely to be only useful when zero weight is essential. Finally, there is no "best" design because the maximum eFOV of 145 degrees is neither very low in volume nor in weight. Remember that these figures are for

particular filtering characteristics Ξ. Code for generating equivalent tables for any Ξ can be found at this project's website [32].

5 EXPERIMENTS AND APPLICATIONS

The ability to provide a wide eFOV for optical convolution allows us to miniaturize previously proposed template-based vision systems. In Fig. 7I, we show our prototype, which consists of a camera (Lu-171, Lumenera Inc.) with custom 3D-printed template assembly. We either cut binary templates into black card paper using a 100-micron laser (VLS3.50, Versa Inc.) or have gray-scale patterns printed on photographic film (PageWorks Inc., http://www.pageworks.com/). We divide the camera photodetector plane into multiple single-template sensor elements using opaque baffles that are created from layered paper to prevent crosstalk between the sensor elements. Snell's window is achieved by attaching laser-cut pieces of acrylic (refractive index $n_1 = 1.5$) to the templates. Ultraviolet-cured optical glue of the same refractive index is used to bind these and fill the air gaps in the templates. Video versions of the results discussed below can be found in [32].

Locating edges. A classical approach to edge detection at a particular scale is to convolve an image with a Laplacian of Gaussian filter [37]. This is often approximated by a

difference-of-Gaussians, and we can do the same here by convolving the scene with two radially symmetric filters in the optical domain. Such a sensor obtains two differently blurred scene measurements, and computes an edge map simply by subtracting corresponding pixels and then thresholding. While the computational savings of this approach are negligible when computing fine scale edges (low-width Gaussians), they increase as the desired edges become more coarse or if the elements are tiled for multiscale edge detection (e.g., [20]).

Fig. 7II demonstrates this using two disk-shaped binary templates of different radii. Like a difference-of-Gaussian operator, differences between corresponding pixels in the two sensor elements produces a band-limited view of the scene (an edge energy map). This is a lensless configuration with two templates having the same heights, $\{d = 0.1 \text{ mm}; u = 3.7 \text{ mm}\}$ and $\{d = 0.2 \text{ mm}; u = 3.7 \text{ mm}\}$ with a (maximized) eFOV of 90 degrees. The figure shows edges of a simple scene with printed words. A naive use of the sensors with suboptimal template height u values of 2 and 5 mm produces incorrect results. Fig. 7III shows an outdoor scene, while Fig. 7IV shows a V-shaped scene viewed by both a simple pinhole and by a wide-FOV Snell's window enhanced sensor, which can "see" more letter edges.

Detecting faces. Traditional face detection can be formulated as a two-step process in which: 1) The image is convolved with a series of templates, and 2) the template responses at each pixel are used as input to a binary classifier. In the past, efficiency has been gained by using "weak" but computationally convenient templates in relatively large numbers [57]. By performing the filtering step optically, we reduce the computational cost further, and because we can use templates with arbitrary spatial patterns and spectral selectivity, we can potentially reduce the number of templates as well.

Optimized spatiospectral templates can surely be learned for discriminating between faces and background, but we leave this for future work. Instead, in Fig. 7V we demonstrate a simple prototype that uses nine *binary templates* learned using a subset of the PubFig Database [34] as positive examples and the method of [23]. The templates are measured in Fig. 7I(c). These are arranged in a lensless configuration $\{d = 0.2 \text{ mm}; u = 5.2 \text{ mm}\}$. While we optimized the design for a 20 degree eFOV, our detector only considers the centers of the nine template responses and does not angularly localize the face. It outputs a response using a linear classifier with no bias term (ensuring invariance to intensity scaling).

Tracking targets. Tracking, in its simplest form, can be implemented as sequential per-frame detection, and thus can be achieved optically using the sensors described above for face detection. If one can afford slightly more computation, then the classifiers used for detection can be combined with a dynamic model to improve performance (e.g., [8], [5]). In either case, we save computation by performing optical filtering-for-matching.

In Fig. 7VI, we show a detector with two templates, a "T" pattern $\{d = 0.2 \text{ mm}; u = 3.7 \text{ mm}\}$ and a small circle $\{d = 0.1 \text{ mm}; u = 3.7 \text{ mm}\}$, optimized for a 90 degree eFOV. After appropriate initialization, we track the target by finding, in a gated region of each subsequent frame, the image point where the pair of template responses is closest

to the initial ones. The nonoptical computation that is required is limited to a small number of subtractions and a minima calculation. We demonstrate tracking for an outdoor scene with obstacles.

## 5.1   Real-Time Tracking with Spectral Templates

The tracking results in the previous discussion were demonstrated on a web-cam platform where power was externally provided and some off-board postprocessing was performed. We next show a proof-of-concept embedded system that performs wide-FOV template-based optical tracking solely with on-board power and computation. Our optical setup is shown in Fig. 8I (*left*). It consists of the two templates in Fig. 7VI embedded in slabs of acrylic. These were laser cut from an acrylic sheet and assembled, by hand, under a microscope. The pieces are held together by Norland optical adhesive that was cured by a Dymax 50AS UV lamp. A section of RoscoLux red filter was cut and attached to the small circular template and this appears reddish.

The embedded platform is an Arduino Pro board, which is a commonly used hobbyist embedded kit [4]. The board processor is an 8-bit ATMega328 16-MHz microcontroller, which is programmed in embedded C. The figure also shows the 5-V power supply for the board, which consists of 3 AA batteries. Implementing convolutions for large image matrices on such a device is prohibitively slow because only 2 KB of SRAM is available at runtime. Demonstrating filtering-based target tracking on such a computationally constrained platform shows how our optical designs, which filter scene radiance off-board, can be advantageous.

We use the Firefly photodetector array from CentEye [11], which is a 128×480 gray-scale imaging sensor with 19.3-micron pixel pitch. While this is a relatively large pixel size, the Firefly has been designed for low-power applications and has a log-response curve between incident radiance and output pixel value. This pixel response allows consistent performance in low-light scenarios that often accompany the use of attenuating templates. The sensor must be calibrated for fixed pattern noise (FPN) by capturing an image of a "blank" scene (such as a white sheet of paper) and storing it within the Arduino's 32 KB of fixed flash memory. The Firefly sensor is fixed on a custom ArduEye Rox1 board from CentEye, that is easily attached to the Arduino.

The ArduEye Rox1 board leaves three binary (0 V (on) or 5 V (off)) output pins free in the Arduino, which we use to drive seven LEDs. We do this with the help of a 74HC595 8-bit shiftout register that converts binary output from the pins into eight states: all LEDs turned off (one state) or each LED turned on individually (seven states). Each of the LEDs indicates the location of the target in the FOV, as shown in the center of Fig. 8I. At the right of (I), we show a frame from a video (available in [32]) of our wide-angle demonstration for tracking a simple red "T" target, displayed on an LCD screen. This demonstration was performed in CVPR 2011 in front of a live audience and in sessions lasting over 4 hours.

In Fig. 8II, we show images from the Arduino system, viewing the same target as in (I). We compare the optical filtering of the sensor with the expected measurements computed in software and show that these are very similar
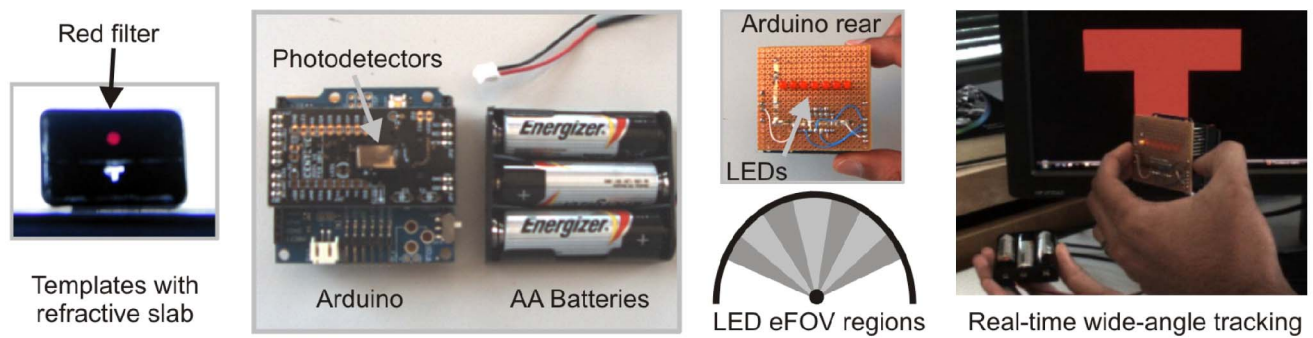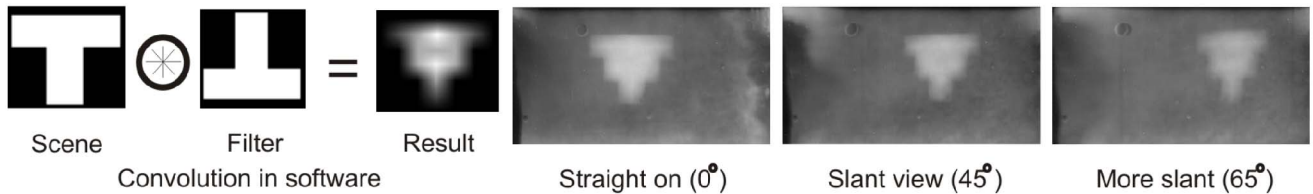
Fig. 8. *Real-time target tracking with an Arduino*. At the left of (I) we show our optics, which consist of templates embedded in a refractive slab. As a proof-of-concept for demonstrating how our optics would include spectral filters, we have placed a red Roscolux filter on one of the templates. We used a custom designed Arduino shield to hold a CentEye Firefly gray-scale sensor. We used a shiftout register to control seven LEDs shown at the rear of the Arduino, which indicate seven different regions in the eFOV. At the right of (I), we show a frame from a video (available in [32]) of our wide-angle tracking demonstration of a simple red "T" target. This demonstration was performed at CVPR 2011 in front of a live audience and over sessions lasting 4 hours. In (II), we compare expected result of filtering, calculated in software, with the sensor measurements from the device. We demonstrate that the responses of the optical filter remain consistent over a wide eFOV, validating the usefulness of the refractive slab.

to each other. In particular, we note that the template response is consistent, even as the angle between the normal to the sensor plane and the sensor-target vector increases. The measured filtered response, although slightly distorted, is consistent even at a 65 degree slant. Therefore, our sensor has a 130 degree eFOV.

5.2 Miniaturized Optics for Fiducial Detection

Fiducials are designed visual features that are artificially placed in an environment to allow easy detection by vision systems. Fiducials are popular in a variety of fields, such as robotics and augmented reality [43], [2], [48], [63], [15], [47], [30]. Recent work has extended these, allowing both active and multispectral fiducials [14], [40], [6]. Locating fiducials can be implemented by applying a large number of filters to a conventionally captured image. In many previous efforts, these implementations have been demonstrated in real time by utilizing the computing power available in a laptop or smartphone. For example, such visual processing is common on quadrotor robots [3].

However, for much smaller classes of air vehicles, the on-board computations required for fiducial detection are too burdensome. Our sensors allow optical filtering that is computationally cheaper and we demonstrate a proof-of-concept device to recognize fiducials on a small, autonomous air vehicle. Our goal here is to demonstrate the usefulness of optical filtering and show the wide-angle capacity of our miniaturized design. For future work, it will be fruitful to consider the design of multispectral templates by extending recent work done on sharing features [49], [53]. However, in this section, as with previous experiments, our optics contain a fixed number of arbitrarily selected binary templates.

In Fig. 9I, we show our miniaturized optics in a sample container, with closeups, under a microscope, from both the top (II) and side (III). We use photolithography and lift-off process for the fabrication. The six binary templates are created by first fabricating a positive optical mask, using a Heidelberg mask writer. This mask is then used to define optical templates on a photoresist coated 150 micron cover glass using a mask aligner. After the exposure and developing process only some photoresist, in the shape of the templates, remains on cover glass. We evaporate 100-nm thick aluminum on the cover glass and soak the glass in acetone. Only the metal deposited on the unexposed photoresist is removed, making the templates transparent while the rest of the cover glass is covered by aluminum, which blocks light.

Polydimethylsiloxane (PDMS) is commonly used in fabrication techniques and is a clear, liquid polymer at room temperature. It can be cured by a variety of methods, after which it becomes a clear solid with a refractive index of about 1.4. We used PDMS for two purposes: first, for making opaque, black baffles that form the bulk of the design in the figure, and second, to embed the templates in a refractive slab that enables a wide eFOV.

Black PDMS sheets for the baffles were created by mixing carbon black particles with clear PDMS. When cured at $65°C$ for 12 hours, this became thin sheets of black, opaque PDMS. The thickness of the black PDMS was controlled by removing layers with scotch tape. Holes in the black PDMS sheet were cut using a VersaLaser and these formed the sensor's baffles. The baffles were placed both above and below the glass slide (carrying the templates), as in (III), by hand, under a microscope.

To create the refractive slab, the entire setup was immersed in clear, liquid PDMS in a vacuum chamber to
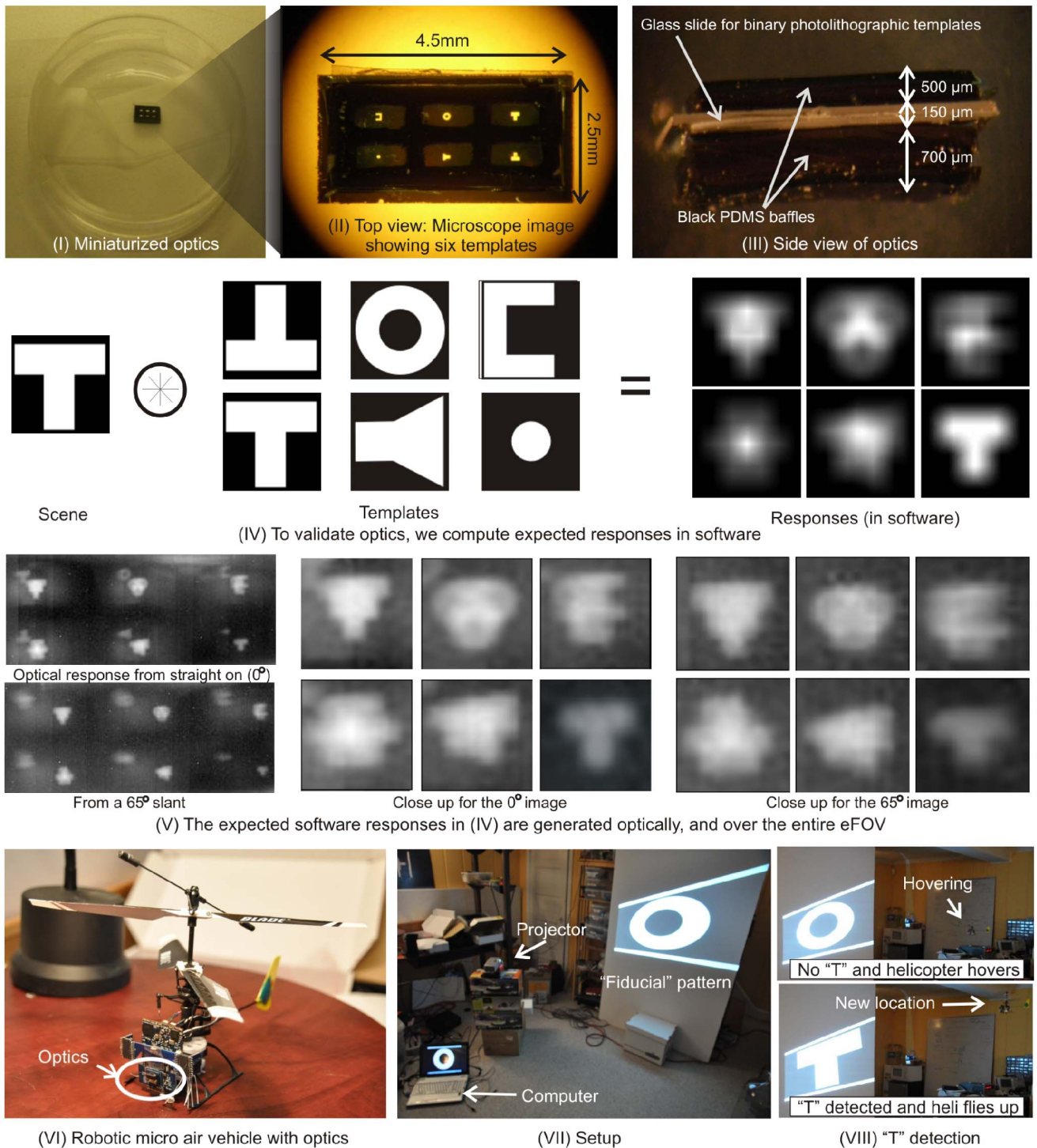
Fig. 9. *Miniaturized optics demonstrated on a micro air vehicle.* In (I) we show our optics in a sample container and also in close-up (II) under a microscope. This is a lensless design with templates embedded in a refractive slab, and its dimensions are shown in (II) and (III). The templates were *arbitrarily selected* and were created by photolithographic techniques with a resolution of 1 micron. In (IV), we show the expected responses of convolution of these templates with a "T" target calculated in software. In (V), we validate our optics by showing that the optical filtering responses are consistent over a wide FOV. In (VI), we show the setup from CentEye of an autonomous micro helicopter, with our optics and our sensor attached. We are able to recognize simple patterns such as the "T" target, and differentiate it from an "O" target and change location based on the type of target. A full video is available at [32].

remove air bubbles. The liquid PDMS was cured at room temperature over 24 hours to form a solid, clear, bubble-free mass around the templates. These miniature templates embedded in PDMS are a version of our lensless design in a refractive slab. (While we did not use them, techniques also exist for fabricating lenslets at this scale [9].) The device was

freed from excess PDMS by slicing by hand with a razor. Imperfect slicing causes the optical surface to be slightly curved. We address this by sandwiching the design between two flat, rectangular pieces of glass with optical glue as an adhesive. This was done in situ and is not shown in the figure.

We visually validated the expected wide FOV behavior of the lensless refractive-slab design in Fig. 9IV. The figure shows the expected (software-simulated) responses of the six arbitrarily-selected templates, to the desired "T" fiducial target. We then measured responses from our optics, placed on a 256×128 Centeye FireFly gray-scale photodetector array, when viewing the same "T" fiducial target on an LCD (Fig. 9V). We captured images from two positions, directly ahead (0 degrees) and at an angle (65 degrees). In both cases, the responses of the template qualitatively match the software responses in Fig. 9IV. We believe the variations that do occur are due to manufacturing errors in our optical design because there are still some steps that occur with human manipulation under a microscope.

In Fig. 9VI, we show a small helicopter robot on which we have placed our sensor, consisting of both our miniaturized optics and the Firefly photodetector array. The helicopter is a converted Blade mcX radio-controlled hobbyist platform that serves as a technology demonstrator for CentEye Inc. A 32-bit Atmel AT32UC3B microcontroller served as the computing power for our detection algorithm, which was implemented in C as template-based matching, and ran on-board. Each of the six subimages (as in Fig. 9V) contributed three values to a feature vector of size 18 that was binarized. This vector was subtracted from the expected responses when viewing a "T" fiducial target, and then the sum of squares of the differences was thresholded.

The algorithm's output is the estimated pixel location of the fiducial's projection. Since we know the template height, we can easily convert the pixel location into an azimuth and elevation angle pair. For future work, we would like to utilize these angles to precisely control the helicopter. Here, we have used the detection of the fiducial to initialize a preset control sequence. This is possible because the helicopter can hover in place, allowing control point-based navigation [7]. We exploit this to move the helicopter by a fixed distance once the fiducial is detected. Visual-based fiducial detection is demonstrated on the air vehicle, and screen shots from our video are shown in Fig. 9VIII.

6 DISCUSSION AND FUTURE WORK

We have described a class of optical designs that allow wide angle filtering of distant scenes. We have demonstrated experiments that validate our theory, and shown a variety of applications. In this section, we will outline some possibilities for future work and provide initial discussions toward these new directions.

SNR analysis. We have explored the space of designs with regards to mass, volume, and eFOV. Extending our work with a formal analysis of noise is also possible. The SNR properties of lenslets could be analyzed with a sensor noise model to formalize the tradeoffs between SNR, volume, mass, and FOV for various designs. Figs. 10a, 10b, and 10c show an example of how the SNR varies over different designs by giving sensor measurements of a face. The template in each of these designs is simply an open aperture. The first is taken with a lensless configuration with a large template height $\{d = 2.5 \text{ mm}; u = 70 \text{ mm}\}$, the second with a reduced template width $\{d = 0.1 \text{ mm}, u = 2.8 \text{ mm}\}$, and the
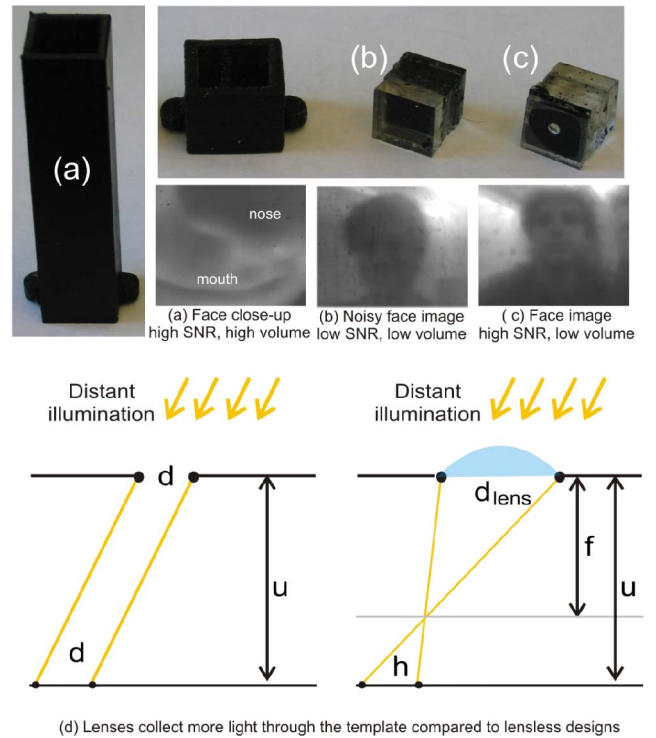

Fig. 10. *SNR issues.* (a) and (b) are pinholes of radii 2.5 and 0.1 mm, while (c) is an lenslet of radii 2.5 mm embedded in acrylic plastic. Lenses collect more light, and hence the SNR advantage of (c) over (b). Additionally, the lenslet (c) allows a compact setup when compared to (a), as shown by the difference in holder size. In (d), we show that just as lenses collect more light for in-focus scenes, they also increase SNR for optical filtering. We compare just the lensless design (b) with the lenslet design (c) for a single, distant scene point. Since we are assuming the scene is infinitely far away, the light rays from this distant scene point are parallel. The two diagrams demonstrate how the information from the scene point is distributed among the photodetectors. We show in the text that $h = d$, that is, information from the distant scene point is distributed among the photodetectors in the same way in these two diagrams. However, since $d_{lens} > d$, the lenslet collects more light and has less noise, for that same distant scene point.

third with an embedded lenslet configuration $\{d = 2.5 \text{ mm}, R = 2.12 \text{ mm}, n_2 = 1.85, \ n_1 = 1.5, u = 12 \text{ mm}\}$. One advantage of lenslets is that the second sensor's volume is smaller than the first, even though the measurement quality appears similar. This fact is illustrated by the difference in size of the optical holders and is related to the analysis presented in this paper. Another seemingly obvious lens advantage is that it collects more light and, hence, the third measurement has better SNR than the second. Fig. 10d shows a diagram where the lensless and lenslet designs are shown viewing a single scene point. Since our scenes are infinitely far away, the light rays from this single scene point are parallel. From similar triangles and the lens equation, we have $h = d_{lens} \frac{u}{v}$, where v is the plane of focus, from the lens equation. Since these two designs have the same eFOV and identical angular supports, from Fig. 4, it is clear that $d = d_{lens} \frac{u}{v}$. Therefore, the lensless width d is equal to h. When $d_{lens} \geq d$ and $u \leq v$, the lenslet collects more light from the scene point and distributes it over the same photodetector area as the lensless design and, therefore, has higher SNR.

Finally, beyond lenslets, the SNR characteristics of the reflective slab are also relevant for any noise analysis done in
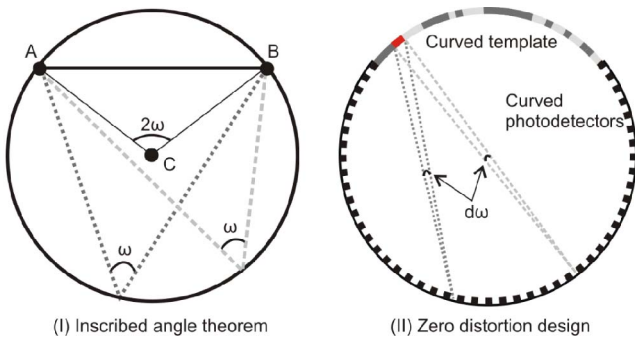
Fig. 11. *Curved sensors*. (I) shows the well-known inscribed angle theorem. In (II), we present a circular curved sensor with a curved template in air (without any refractive slab). The angular support of the template as well as the angular support of each printed "dot" on the template is identical for each photodetector. Such a sensor would have zero distortion and would allow for perfect optical filtering over a 180 degree FOV.

the future. For example, fresnel reflection occurs for dielectric surfaces, such as the acrylic slabs used in our experiments. According to Fresnel's law, all light incident at a grazing angle will be reflected, and we will never achieve full 180 degree eFOV. More importantly for the SNR analysis, the percentage of light that is reflected increases as we approach the grazing angle, reducing the measured signal.

Curved sensors and templates. This work is only one example of how optical processing can help achieve vision on a tight budget. We may consider using other optical elements, such as adaptive templates [41], artificial insect eyes [28], [27], and multiplexed configurations [55], as and when they become widely available in small, low-power form factors. In particular, curved sensors [31], [33], [17] are increasingly becoming a reality. In Fig. 11, we show a possible curved sensor design for optical filtering over a wide FOV. We note, in (I), that the inscribed angle theorem, which is well known in geometry, states that the angle subtended by an arc and its corresponding chord on any part of the circle is identical to half the angle subtended at the center.

We propose a sensor with a circular array of photodetectors and a curved template in air (not a refractive slab), as in Fig. 11II, which takes advantage of this property, for any desired angular support ω. This sensor has a curved template, which lies along the same circle as the photodetectors. Every printed "dot" on the circular template arc also follows the inscribed angle theorem and has fixed angular support across
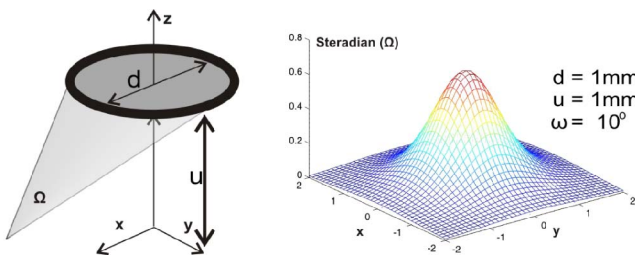

Fig. 12. *Solid angle*. Left: For a circular support of a template, the solid angle of a lensless design Ω can easily be calculated from the angular support of its 2D design ω from the equation for the solid angle of a cone, $2\pi(1 - \cos \omega)$. Right: An illustration of the solid angle for particular design values. An identical equation for Ω follows for lenslets in air from the discussion in Fig. 4.
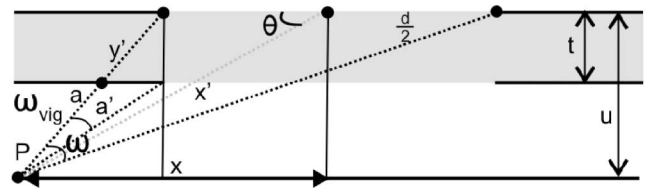

Fig. 13. *Aperture vignetting*.

the photodetectors. Therefore, we obtain zero distortion in the angular dot pitch $d\omega$, which we defined previously. *This circular design would be a "perfect" optical filtering sensor with zero distortion and 180 degree eFOV.*

Extensions to 3D. In Fig. 12, we discuss a 3D analysis of the lensless design in air. Note that although the support of the template is circular, the distribution of gray-scale values within this region can be any pattern. Given a 2D version of our design, with angular support ω (1), we can compute the 3D angular support Ω from the equation for the solid angle of a vertex of a cone $\Omega = 2\pi(1 - \cos \omega)$. We provide an illustration of how Ω varies for particular lensless parameters in the figure. It also clear from Fig. 4 that the solid angle of a 3D lenslet in air would utilize the same cone equation. Therefore, we have provided a way to find the solid angle for the 3D lensless and lenslet cases in air, from our 2D equations involving the solid angle ω.

Since our designs are symmetric, their weight and volume in 3D follow monotonically from the 2D analysis: If two designs are such that one is heavier in our 2D analysis, then they must have a similar relationship in 3D. Finally, we note two directions here for future work. The first is to understand how the user-defined tolerance Δ changes in 3D. The second is to find equations for the solid angle of a refractive slab in 3D.

Learning the best spatio-spectral templates. Our analysis of optical designs assumes that a set of templates have been pre-chosen or prelearned for the task at hand. Developing machine learning tools specifically for our platforms may be a worthwhile direction to pursue. These tools should account for fabrication constraints (for example, resolution and bit depth) and template distortion (Δ) during learning, and they should be capable of producing templates that not only have discriminative spatial patterns, but also have discriminative spectral responses as well. Indeed, the ability to easily specify a unique spectral profile (over UV, VIS, and NIR) for each template in our sensors may enhance their utility by endowing them with characteristics, such as lighting, pose, and scale insensitivity, typically associated with conventional vision systems.

The effect of aperture thickness. We explain the effect of aperture thickness in Fig. 13. This effect can be included in our designs simply by subtracting the obstructed "vignetted" solid angle ω_{vig} from the angular support ω for a particular design. Total vignetting occurs when $\arctan(\frac{t}{d}) = \arctan(\frac{u-t}{x-\frac{d}{2}})$. No vignetting occurs when $\frac{-d}{2} \geq x \leq \frac{d}{2}$. Elsewhere, the angular support decreases by

$$\omega_{vig} = \arccos \sqrt{\frac{(y' + a)^2 + (a')^2 - t^2}{2(y' + a)(a')}},$$

where

$$y' = \sqrt{\left(\frac{t^2\left(x - \frac{d}{2}\right)^2 + u^2 t^2}{u^2}\right)},$$

$$a = \sqrt{(u-t)^2 + \left(x - \frac{t\left(x - \frac{d}{2}\right)}{u} - \frac{d}{2}\right)^2}, \quad \text{and}$$

$$a' = \sqrt{(u-t)^2 + \left(x - \frac{d}{2}\right)^2}.$$

ACKNOWLEDGMENTS

REFERENCES

[1] *Zeemax Optical Software,* http://www.zemax.com/, 2010.
[2] F. Ababsa and M. Mallem, "A Robust Circular Fiducial Detection Technique and Real-Time 3D Camera Tracking," *J. Multimedia,* vol. 3, pp. 34-41, 2008.
[3] E. Altug, J. Ostrowski, and R. Mahony, "Control of a Quadrotor Helicopter Using Visual Feedback," *Proc. IEEE Int'l Conf. Robotics Automation,* 2002.
[4] Arduino and Pro, "Arduino Website," http://www.arduino.cc/, 2012.
[5] S. Avidan, "Support Vector Tracking," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2001.
[6] H. Bagherinia and R. Manduchi, "A Theory of Color Barcodes," *Proc. IEEE Color and Photometry in Computer Vision Workshop,* 2011.
[7] G. Barrows, J. Chahl, and Y. Srinivasan, "Biomimetic Visual Sensing and Flight Control," *Proc. Bristol UAV Conf.,* 2002.
[8] M.J. Black and A.D. Jepson, "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation," *Int'l J. Computer Vision,* vol. 26, pp. 63-84, 1998.
[9] N. Borrelli, *Microoptics Technology: Fabrication and Applications of Lens Arrays and Devices.* CRC Press, 1999.
[10] V. Brajovic and T. Kanade, "Computational Sensor for Visual Tracking with Attention," *IEEE J. Solid State Circuits,* vol. 33, no. 8, pp. 1199-1207, Aug. 1998.
[11] CentEye and Inc, "Centeye Website," http://www.centeye.com/, 2012.
[12] A. Chandrakasan, N. Verma, J. Kwong, D. Daly, N. Ickes, D. Finchelstein, and B. Calhoun, "Micropower Wireless Sensors," *Proc. NSTI Nanotechnology Conf.,* 2006.
[13] V. Chari and P. Sturm, "Multi-View Geometry of the Refractive Plane," *Proc. British Machine Vision Conf.,* 2009.
[14] Y. Cho and U. Neumann, "Multi-Ring Color Fiducial Systems for Scalable Fiducial Tracking Augmented Reality," *Proc. IEEE Virtual Reality Ann. Int'l Symp.,* 1998.
[15] D. Claus and A. Fitzgibbon, "Visual Marker Detection and Decoding in AR Systems: A Comparative Study," *Proc. Int'l Symp. Mixed Augmented Reality,* 2002.
[16] Collection, *Flying Insects and Robots.* Springer, 2009.
[17] O. Cossair, D. Miau, and S. Nayar, "A Scaling Law for Computational Imaging with Spherical Optics," *J. Optical Soc. Am. A,* vol. 28, pp. 2540-2553, 2011.
[18] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-Pixel Imaging via Compressive Sampling," *IEEE Signal Processing Magazine,* vol. 25, no. 2, pp. 83-91, Mar. 2008.
[19] M. Edge and I. Turner, *The Underwater Photographer.* Focal Press, 1999.
[20] J.H. Elder and S.W. Zucker, "Local Scale Control for Edge Detection and Blur Estimation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 7, pp. 699-716, July 1998.
[21] C. Farabet, C. Poulet, and Y. LeCun, "An FPGA-Based Stream Processor for Embedded Real-Time Vision with Convolutional Networks," *Proc. Fifth IEEE Workshop Embedded Computer Vision,* 2009.
[22] A. Fluegel http://glassproperties.com/, 2007.
[23] I. Gkioulekas and T. Zickler, "Dimensionality Reduction Using the Sparse Linear Model," *Proc. Neural Information Processing Systems,* 2011.
[24] J.W. Goodman, *Introduction to Fourier Optics.* McGraw-Hill, 1968.
[25] B. Gyselinckx, C. Van Hoof, J. Ryckaert, R. Yazicioglu, P. Fiorini, and V. Leonov, "Human++: Autonomous Wireless Sensors for Body Area Networks," *Proc. IEEE Custom Integrated Circuits Conf.,* pp. 13-19, 2006.
[26] H.P. Herzig, *Micro-Optics: Elements, Systems and Applications.* Taylor and Francis, 1999.
[27] S. Hiura, A. Mohan, and R. Raskar, "Krill-Eye: Superposition Compound Eye for Wide-Angle Imaging Via Grin Lenses," *Proc. Workshop Omni Directional Vision,* 2009.
[28] K. Jeong, J. Kim, and L. Lee, "Biologically Inspired Artificial Compound Eyes," *Science,* vol. 312, pp. 557-561, 2006.
[29] M. Karpelson, G. Wei, and R.J. Wood, "Milligram-Scale High-Voltage Power Electronics for Piezoelectric Microrobots," *Proc. IEEE Int'l Conf. Robotics Automation,* 2009.
[30] H. Kato and M. Billinghurst, "Marker Tracking and hmd Calibration for a Video-Based Augmented Reality Conferencing System," *Proc. Second IEEE ACM Int'l Workshop Augmented Reality,* 1999.
[31] H. Ko, G. Shin, S. Wang, M. Stoykovich, J. Lee, D. Kim, J. Ha, Y. Huang, K. Hwang, and J. Rogers, "Curvilinear Electronics Formed Using Silicon Membrane Circuits and Elastomeric Transfer Elements," *Small,* vol. 5, pp. 2703-2709, 2009.
[32] S.J. Koppal, "Toward Micro Vision Sensors Website," http://www.koppal.com/microvisionsensors.html, 2012.
[33] G. Krishnan and S.K. Nayar, "Towards a True Spherical Camera," *Proc. SPIE,* 2009.
[34] N. Kumar, A.C. Berg, P. Belhumeur, and S.K. Nayar, "Attribute and Simile Classifiers for Face Verification," *Proc. 12th IEEE Int'l Conf. Computer Vision,* 2009.
[35] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient Based Learning Applied to Document Recognition," *Proc. IEEE,* vol. 86, no. 11, pp. 2278-2324, Nov. 1998.
[36] A. Levin, R. Fergus, and B. Freeman, "Image and Depth from a Conventional Camera with a Coded Aperture," *Proc. ACM Siggraph,* 2007.
[37] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. Royal Soc. London,* vol. 207, pp. 187-217, 1980.
[38] K. Mielenz, "On the Diffraction Limit for Lensless Imaging," *J. Research of the NIST,* vol. 104, pp. 479-485, 1999.
[39] K. Miyamoto, "Fish Eye Lens," *J. Optical Soc. Am.,* vol. 54, pp. 1060-1061, 1964.
[40] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar, "Bokode: Imperceptible Visual Tags for Camera Based Interaction from a Distance," *Proc. ACM Siggraph,* 2009.
[41] S.K. Nayar, V. Branzoi, and T.E. Boult, "Programmable Imaging: Towards a Flexible Camera," *Int'l J. Computer Vision,* vol. 70, pp. 7-22, 2006.
[42] R. Ng, "Fourier Slice Photography," *ACM Trans. Graphics,* vol. 24, pp. 735-744, 2005.
[43] E. Olson, "AprilTag: A Robust and Flexible Visual Fiducial System," *Proc. IEEE Int'l Conf. Robotics and Automation,* 2011.
[44] M. O'Toole and K. Kutulakos, "Optical Computing for Fast Light Transport Analysis," *Pro. ACM Siggraph Asia,* 2010.
[45] F. Pedrotti and L. Pedrotti, *Introduction to Optics.* Benjamin Cummings, 2006.
[46] C. Raghavendra, K. Sivalingam, and T. Znati, *Wireless Sensor Networks.* Springer, 2004.

[47] J. Rekimoto and Y. Ayatsuka, "CyberCode: Designing Augmented Reality Environments with Visual Tags," *Proc. ACM Designing Augmented Reality Environments,* 2000.

[48] J. Sattar, E. Bourque, P. Giguere, and G. Dudek, "Fourier Tags: Smoothly Degradable Fiducial Markers for Use in Human-Robot Interaction," *Proc. Fourth Canadian Conf. Computer and Robot Vision,* 2007.

[49] S. Shalev-Shwartz, Y. Wexler, and A. Shashua, "Shareboost: Efficient Multiclass Learning with Feature Sharing," *Proc. Neural Information Processing Systems,* 2011.

[50] E. Steltz and R. Fearing, "Dynamometer Power Output Measurements of Miniature Piezoelectric Actuators," *IEEE/ACM Trans. Mechatronics,* vol. 14, no. 1, pp. 1-10, Feb. 2009.

[51] R. Swaminathan, M. Grossberg, and S. Nayar, "Caustics of Catadioptric Cameras," *Proc. Eighth IEEE Int'l Conf. Computer Vision,* 2001.

[52] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka, "Thin Observation Module by Bound Optics (TOMBO): Concept and Experimental Verification," *Applied Optics,* vol. 40, pp. 1806-1813, 2001.

[53] A. Torralba, K. Murphy, and W. Freeman, "Sharing Visual Features for Multiclass and Multiview Object Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 5, pp. 854-869, May 2007.

[54] T. Treibitz, Y. Schechner, and H. Singh, "Flat Refractive Geometry," *Proc. IEEE Conf. Computer Vision Pattern Recognition,* 2008.

[55] S. Uttam, N. Goodman, M. Neifeld, C. Kim, R. John, J. Kim, and D. Brady, "Optically Multiplexed Imaging with Superposition Space Tracking," *Optics Express,* vol. 17, pp. 1691-1713, 2009.

[56] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing," *Proc. ACM Siggraph,* 2007.

[57] P.A. Viola and M.J. Jones, "Robust Real-Time Face Detection," *Int'l J. Computer Vision,* vol. 57, pp. 137-154, 2004.

[58] R. Volker, M. Eisner, and K. Weible, "Miniaturized Imaging Systems," *Microelectronic Eng.,* vol. 67/68, pp. 461-472, 2003.

[59] A. Wilhelm, B. Surgenor, and J. Pharoah, "Evaluation of a Micro Fuel Cell as Applied to a Mobile Robot," *Proc. IEEE Int'l Conf. Mechatronics Automation,* 2005.

[60] W. Wolf, B. Ozer, and T. Lv, "Smart Cameras as Embedded Systems," *Computer,* vol. 35, no. 9, pp. 48-53, Sept. 2002.

[61] R.W. Wood, *Physical Optics.* Macmillan, 1911.

[62] F. Yu and S. Jutamulia, *Optical Pattern Recognition.* Cambridge Univ. Press, 1998.

[63] X. Zhang, S. Fronz, and N. Navab, "Visual Marker Detection and Decoding in AR Systems: A Comparative Study," *Proc. First Int'l Symp. Mixed and Augmented Reality,* 2002.

[64] A. Zomet and S. Nayar, "Lensless Imaging with a Controllable Aperture," *Proc. IEEE Conf. Computer Vision Pattern Recognition,* 2006.

Sanjeev J. Koppal received the BS degree from the University of Southern California in 2003, and the master's and PhD degrees from the Robotics Institute at Carnegie Mellon University. He is currently a postdoctoral associate at Harvard University. His interests span computer vision and computational photography and include novel sensors, digital cinematography, 3D cinema, light-field rendering, appearance modeling, 3D reconstruction, physics-based vision, and active illumination. He is a member of the IEEE.

Ioannis Gkioulekas received degrees in electrical and computer engineering from the National Technical University of Athens, Greece, and is currently working toward the PhD degree in electrical engineering at the Harvard School of Engineering and Applied Sciences, where he is a member of the Graphics, Vision and Interaction Group. He is a student member of the IEEE.
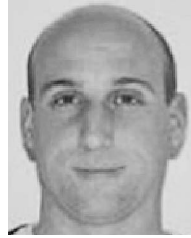
Travis Young graduated from the University of Maryland, College Park, in 2007 with a degree in electrical engineering, and he is currently working toward the master's degree there. Since 2008, he has been with Centeye, Inc., designing minimalist vision systems for micro air vehicles to aid in autonomous navigation. He is a member of the IEEE.

Hyunsung Park received the BS and MS degrees in electrical engineering from Seoul National University, Seoul, Korea, in 2006 and 2008, respectively, and is currently working toward the PhD degree in electrical engineering at Harvard University, Cambridge, Massachusetts. His PhD research topic is vertical nanowire-based optical filters and photodetectors. He is a student member of the IEEE.

Geoffrey L. Barrows received the BS degree in applied mathematics from the University of Virginia, the MS degree in electrical engineering from Stanford University, and the PhD degree in electrical engineering from the University of Maryland at College Park. He is the founder of Centeye, a company that specializes in the development of insect vision for robotics. In 2003, he was recognized as a "young innovator" by being included in the MIT Technology Review's TR100 list. He is a member of the IEEE.

Kenneth B. Crozier received the undergraduate degrees in electrical engineering (first class honors, with medal) and physics from the University of Melbourne, Australia, and the PhD degree in electrical engineering from Stanford University in 2003. He is an associate professor of electrical engineering at Harvard University. His research interests include nanooptics, with an emphasis on plasmonics for optical manipulation and surface enhanced Raman spectroscopy. He received the US National Science Foundation (NSF) CAREER Award in 2008. He is a member of the IEEE.

Todd Zickler received the PhD degree in electrical engineering from Yale University in 2004. He is the Gordon McKay Professor of Electrical Engineering and Computer Science in the School of Engineering and Applied Sciences at Harvard University. He is the director of the Harvard Computer Vision Laboratory and a member of the Graphics, Vision and Interaction Group. His research is focused on modeling the interaction between light and materials, and developing systems to extract scene information from visual data. His work is motivated by applications in face, object, and scene recognition; image-based rendering; image retrieval; image and video compression; robotics; and human-computer interfaces. He received the US National Science Foundation (NSF) Career Award and a Research Fellowship from the Alfred P. Sloan Foundation. His research is funded by the NSF, the US Army Research Office, and the US Office of Naval Research. He is a member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.