This CVPR2015 paper is the Open Access version, provided by the Computer Vision Foundation.
The authoritative version of this paper is available in IEEE Xplore.

Privacy Preserving Optics for Miniature Vision Sensors

Francesco Pittaluga and Sanjeev J. Koppal
University of Florida, Electrical and Computer Engineering Dept.
216 Larsen Hall Gainesville, FL 32611-6200
f.pittaluga@ufl.edu and sjkoppal@ece.ufl.edu

Abstract

The next wave of micro and nano devices will create a world with trillions of small networked cameras. This will lead to increased concerns about privacy and security. Most privacy preserving algorithms for computer vision are applied after image/video data has been captured. We propose to use privacy preserving optics that filter or block sensitive information directly from the incident light-field before *sensor measurements are made, adding a new layer of privacy. In addition to balancing the privacy and utility of the captured data, we address trade-offs unique to miniature vision sensors, such as achieving high-quality field-of-view and resolution within the constraints of mass and volume. Our privacy preserving optics enable applications such as depth sensing, full-body motion tracking, people counting, blob detection and privacy preserving face recognition. While we demonstrate applications on macro-scale devices (smartphones, webcams, etc.) our theory has impact for smaller devices.*

1. Introduction

Our world is bursting with ubiquitous, networked sensors. Even so, a new wave of sensing that dwarfs current sensor networks is on the horizon. These are miniature platforms, with feature sizes less than 1mm, that will appear in micro air vehicle swarms, intelligent environments, body and geographical area networks. Equipping these platforms with computer vision capabilities could impact security, search and rescue, agriculture, environmental monitoring, exploration, health, energy, and more.

Yet, achieving computer vision at extremely small scales still faces two challenges. First, the power and mass constraints are so severe that full-resolution imaging, along with post-capture processing with convolutions, matrix inversions, and the like, are simply too restrictive. Second, the privacy implications of releasing trillions of networked, tiny cameras into the world would mean that there would likely be significant societal pushback and legal restrictions.

In this paper, we propose a new framework to achieve both power efficiency and privacy preservation for vision on small devices. We build novel optical designs that filter incident illumination from the scene, before image capture. This allows us to attenuate sensitive information while capturing exactly the portion of the signal that is relevant to a particular vision task. In this sense, we seek to generalize the idea of privacy preserving optics beyond specialized efforts (cylindrical lenses [45], thermal motion sensors [7]). We demonstrate privacy preserving optics that enable accurate depth sensing, full-body motion tracking, multiple people tracking, blob detection and face recognition.

Our optical designs filter light before image capture and represent a new axis of privacy vision research that complements existing "post image capture" hardware and software based approaches to privacy preservation, such as de-identification and cryptography. Like these other approaches, we seek both data-utility and privacy protection in our designs. Additionally, for miniature sensors, we must also balance the performance and privacy guarantees of the system with sensor characteristics such as mass/volume, field-of-view and resolution. In this paper, we demonstrate applications on macro-scale devices (smartphones, webcams, etc.), but our theory has impact for smaller devices.

Our contributions are

1. To our knowledge, we are the first to demonstrate k-anonymity preserving optical designs for faces. We also provide theory to miniaturize these designs within the smallest sensor volume.
2. We show how to select a defocus blur that provides a certain level of privacy over a working region, within the limits of sensor size. We show applications where defocus blur provides both privacy and utility for time-of-flight and thermal sensors.
3. We implement scale space analysis using an optical array, with most of the power hungry difference-of-gaussian computations performed pre-capture. We demonstrate human head tracking with this sensor. We provide an optical version of the knapsack problem to miniaturize such multi-aperture optical privacy preserving sensors in the smallest mass/volume.

1.1. Related Work

Applied optics and computational photography for privacy preserving computer vision. [7] proposed a system using thermal motion sensors that enables two-person motion tracking in a room. [45] used a line sensor and cylindrical lens to detect a person's position and movement. [53] controlled the light-transport to shadow sensitive regions, removing data-utility in those areas. Our proposed optical systems offer significant improvement over these systems in terms of data-utility by capturing appropriately modulated two dimensional sensor readings.

Privacy preserving computer vision algorithms. Pixelation, Gaussian blurring, face swapping [4] and black-out [5], provide privacy by trading off image utility [47, 41]. More complex encryption based schemes [64, 17, 39], enable recovery of the original data via a key. Other non-distortion based methods based on k-anonymity [63], provably bound face recognition rate while maintaining image utility [48, 29, 28, 15, 2]. We demonstrate the advantages of performing some of these algorithms (such as k-anonymity and defocus blur) in optics, prior to image capture.

Optics-based cryptography. [32] proposed an optics-based encrypted communication framework where, for example, random cryptographic bits are kept safe by volumetric scattering materials. Our work exploits optics to construct privacy preserving sensors that process illumination directly from scenes.

Embedded systems and privacy preserving computer vision. The embedded vision community has proposed a number of privacy sensors [44, 11, 69] which transform the vision data at the camera level itself or offline and then use encryption or other methods to manage the information pipeline. The hardware integration decreases such systems' susceptibility to attacks. Our privacy preserving optics provide another complementary layer of security by removing sensitive data before image capture through optical "off-board" processing. Further, our optical knapsack approach is a miniature analog to larger camera sensor network coverage optimizations [21, 19, 60, 20].

Efficient hardware for small-scale computer vision. The embedded systems community has proposed many vision techniques for low-power hardware [70, 6, 37]. That said, for micro-scale platforms, the average power consumption is often in the range of milli-Watts or micro-Watts [31, 10, 8, 59, 61, 68]. In these scenarios, our approach of jointly considering optics, sensing, and computation within the context of platform constraints will be crucial.

Face De-blurring. Despite significant advances in image and video de-blurring [54, 72, 51, 50, 24, 3, 14], de-blurring heavily blurred images is still an open problem. In this paper, some designs that use optical defocus for privacy may be susceptible to reverse engineering.

Filtering in applied optics and computational photography. Fourier optics [27, 71] has limited impact for miniature vision systems that must process incoherent scene radiance. However, controllable PSFs in conjunction with post-capture processing are widely used in computer vision [57, 49, 38, 25]. In contrast to these approaches, we seek optics like [34, 35, 74, 46] that distill the incoming light-field for vision applications.

Compressive Sensing CS techniques have found application in imaging and vision [66, 16] and some approaches use random optical projection [16], which could be augmented with privacy preserving capabilities. Further, optical projection and classification have been integrated (without any privacy preservation) as in [13]. Some of these algorithms are linear [73, 1, 65, 12] and, in future work, we may consider implementing these within our optical frame work.

2. Single Aperture Privacy Preserving Optics

We introduce two optical designs that perform privacy preserving computations on the incident light-field *before* capture. The first design, performs optical averaging and enables k-anonymity image capture. The second, uses an aperture mask to perform angular convolutions and enables privacy enhancing image blur. For each design we describe how to trade-off the optics' mass/volume with sensor characteristics such as resolution and field-of-view (FOV).

2.1. Optical K-Anonymity

K-anonymity for faces [63, 48] enables face de-identification by averaging together a target face image with $k-1$ of its neighbors (according to some similarity metric). The resulting average image has an algorithm-invariant face recognition rate bound of $\frac{1}{k}$. We present what is, to our knowledge, the first ever optical implementation of k-anonymity for faces. Our system, illustrated in Fig. 1(I), consists of a sensor (approximated by an ideal pinhole camera) whose viewing path is split between the scene and an active optical mask, such as a projector or electronic display. The irradiance I measured at each sensor pixel (x, y) that views a scene point P is given by,

$$I(x,y) = e_P\, I_P + e_M \sum_{1 \le i \le k-1} I_{mask}(w_i F_i(H(x,y))), \tag{1}$$

where I_P is the radiance from P, F_i are digital images of the $k-1$ nearest neighbors, I_{mask} maps a mask pixel intensity to its displayed radiance, w_i are user defined weights and H is a transformation between the sensor and mask planes. e_P and e_M are the ratios of the optical path split between the scene and the mask, and these can range from 0 to 1. We use planar non-polarizing half-mirrors in Fig. 1, so $e_P = e_M = 0.5$ and the sensor exposure must be doubled to create full intensity k-anonymized images.

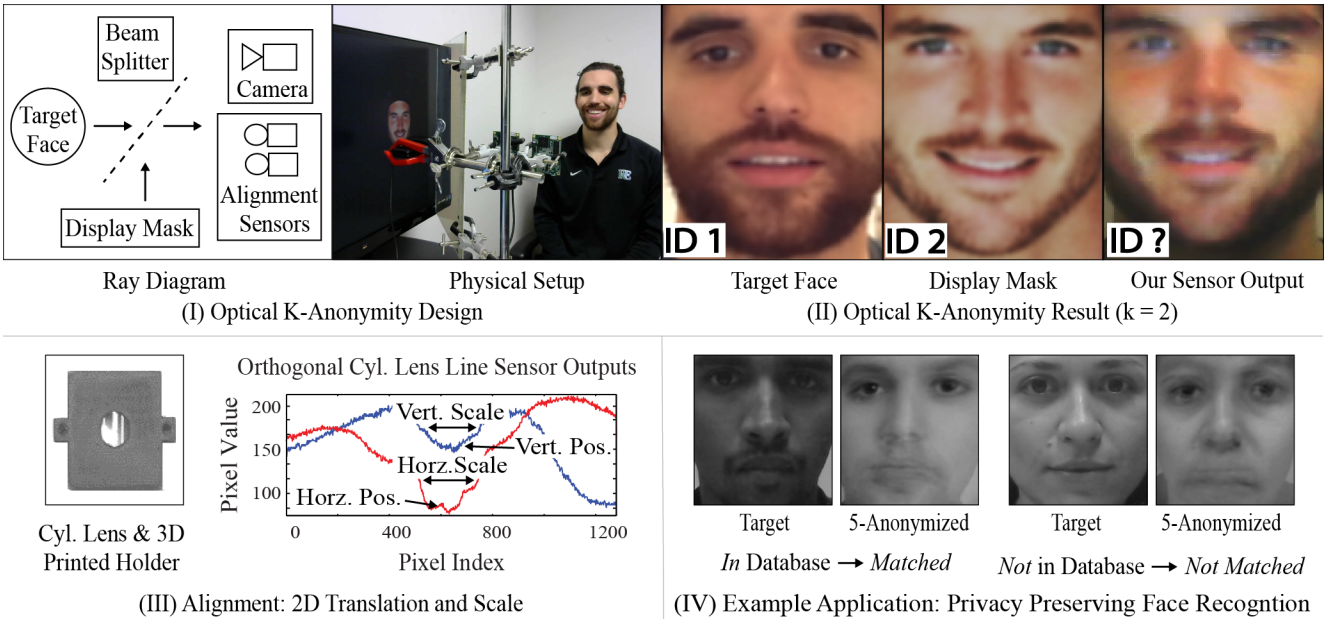


Figure 1. **Optical K-Anonymity for Faces**. Here, we show our design and results for, to our knowledge, the first ever optics-based implementation of k-anonymity for faces [48]. In (I) we show the ray diagram and physical setup for our design whose primary input is k, the number of faces to anonymize a target face with. Light from a real target face is merged via a beamsplitter with illumination from a display showing the $k-1$ nearest neighbors and captured by a conventional sensor. The output is a k-anonymized face, directly captured by our sensor, as shown in (II). Finding the $k-1$ neighbors and 2D translation/scaling alignment, between the target face and the $k-1$ displayed faces, is achieved using two orthogonally-oriented line sensors with cylindrical lenses (III). The scale and position of the target face is found by identifying local extrema of the intensity profiles. Lastly, in (IV) we show an example application that enables privacy preserving face recognition for individuals in a membership class and maintains anonymity for individuals outside of the membership class.

Our implementation in Fig. 1 uses an LED, a webcam, a beam splitter, and two line sensors with orthogonally-oriented 6mm focal length cylindrical lenses. The output is a k-anonymized face, directly captured, at 30 FPS, by our sensor, as shown in Fig. 1(II). Finding the $k-1$ neighbors and 2D translation/scaling alignment, between the target face and the $k-1$ displayed faces, is achieved using the two line sensors with cylindrical lenses, which have been shown to be privacy preserving [45]. The scale and position of the target face is found by identifying local extrema of the intensity profiles as shown in Fig. 1(III). The linear combination of the k-1 faces displayed by the LCD is generated by aligning the k-1 faces, with any alignment method [4, 9], and computing an appropriately weighted sum of the k-1 faces.

Discussion: The use of a display commits the system to continuous power use which makes miniaturization difficult. However, in the next section we discuss how to reduce the volume of the optics for small form factor platforms. In addition, we have assumed the $k-1$ neighbors F_i in Eq. 1 are captured under similar illumination environments to the target face. In the future, we will relax this by using an additional single photodetector element, which is also privacy preserving as it only captures a single intensity value, to set the linear weights w_i in Eq. 1 to compensate for the image intensity differences. Additionally, the display is susceptible to physical tampering that might prevent k-anonymity. Finally, in the current implementation, access to the database could allow an adversary to remove k-anonymity. In future implementations we plan to randomize the value k, the choice of k neighbors and the blending weights w_i to make de-anonymity combinatorially hard.

2.1.1 Miniaturizing K-Anonymity Optics

Optical k-anonymity requires that the resolution of the display be equal to or greater than the resolution of the sensor. Here we discuss how to reduce the size of the k-anonymity optical setup while still maintaining the desired display resolution. We assume that the camera sensor in Fig. 1 is optimally miniaturized by a method such as [34]. For clarity we consider a 2D ray diagram, but since our optics are symmetric these arguments hold in three dimensions. Let the beamsplitter angle be fixed at ϕ and the sensor FOV be θ. Let the minimum size of the mask that still affords the desired resolution be M_{min}. W.l.o.g let the mask be perpendicular to the reflected optical axis.

This leaves just two degrees of freedom for the k-anonymity optics; the sensor-beamsplitter distance l_{beam}

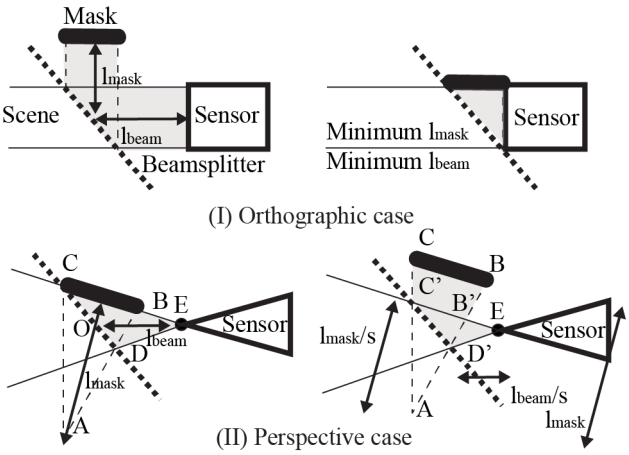


Figure 2. **Miniaturizing Optical K-same** We demonstrate how to reduce the volume occupied by the display and beamsplitter, determined by l_{beam} and l_{mask}. For the perspective case, we show that there exists two configurations with identical, minimum volume.

along the sensor's optical axis and the mask-beamsplitter distance l_{mask} along the reflected optical axis. In an orthographic version of k-anonymity optics shown in Fig. 2 (I), the size of the mask does not change as it is translated towards the sensor. Therefore, a mask of minimum size M_{min} can be moved as close as possible to the sensor without occluding the field-of-view as in Fig. 2 (I).

In the perspective case [26] the size of the mask reduces as it slides along the pencil of rays, as in Fig. 2 (II). Once the minimum mask size M_{min} is reached, that configuration has the minimum optical size, given by $\triangle CDE$'s area.

We show that there exists an alternate choice, in the perspective case, for the minimum optical size. To maintain the minimum resolution, any mask position closer to the sensor must be vertically shifted, as in Fig. 2 (II). The area of these optics is given by $\triangle C^{'}D^{'}E + C^{'}B^{'}BC$. From similar triangles, we can write $\triangle C^{'}D^{'}E$ as being created from $\triangle CDE$ by a scale factor $\frac{1}{s}$, and then equate the two configurations in Fig. 2 (II),

$$\triangle CDE(1 - \frac{1}{s}) = C^{'}B^{'}BC. \quad (2)$$

Consider $\triangle CDE = \triangle COE + \triangle ODE$. From the angle-side-angle theorem, this becomes,

$$\triangle CDE = \frac{l_{beam}^2 \sin\frac{\theta}{2} \sin\phi}{2 \sin(\frac{\theta}{2} - \phi)} + \frac{l_{beam}^2 \sin\frac{\theta}{2} \sin\phi}{2 \sin(\frac{\theta}{2} + \phi)}. \quad (3)$$

Since $\triangle AB^{'}C^{'}$ is a scaled version of $\triangle ABC$, the quadrilateral area $C^{'}B^{'}BC =$

$$\triangle ABC(1 - \frac{1}{s^2}) = \frac{M_{min} l_{mask}}{2}(1 - \frac{1}{s^2}). \quad (4)$$

Putting Eq. 3 and Eq. 4 into Eq. 2, and setting constant $C_1 = \frac{\sin\frac{\theta}{2} \sin\phi}{2 \sin(\frac{\theta}{2} - \phi)} + \frac{\sin\frac{\theta}{2} \sin\phi}{2 \sin(\frac{\theta}{2} + \phi)}$,

$$s = \frac{M_{min} l_{mask}}{2C_1 l_{beam}^2 - M_{min} l_{mask}}, \quad (5)$$

which is an equation for the scaling factor s such that the two designs in Fig. 2 (II) have the same area. Therefore we have found two designs that provide the required resolution within the smallest optical dimensions.

Example Application: Privacy Preserving Face Recognition: Recent efforts have resulted in privacy preserving face recognition frameworks [58, 22, 52, 33]. Here we show a similar example application, using optical k-same, that allows recognition of membership to a class while preserving privacy. Each target is first anonymized via optical k-same with k-1 faces corresponding to individuals that are not in the membership class and are not known to the party performing face recognition. The anonymized face is compared to each face in the membership class using a similarity metric. If the similarity score is greater than a threshold then the anonymized face is matched with that individual. With no match, the system returns the k-anonymized face.

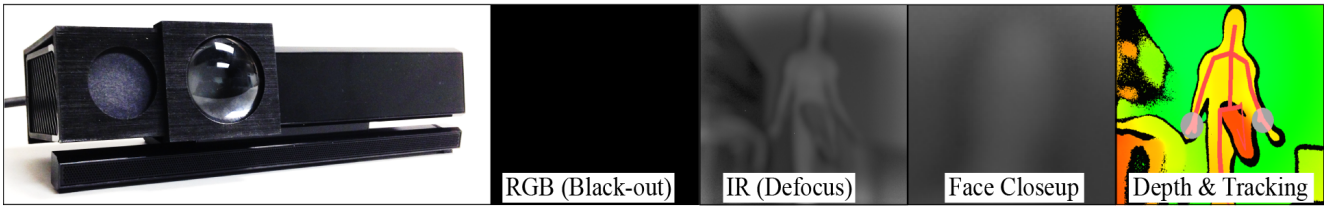
We simulated this system using two subsets of the FERET Database [55], each containing a single image of a set of people (See supplementary document at [56]). For k = $\{2, 4, 6, 8, 10\}$, 100 individuals from one subset were randomly selected as targets and anonymized with their $k-1$ nearest neighbors found in the same subset by simulating the effect of the cylindrical lens by integrating the image vertically and matching with the cosine similarity. The similarity between this k-anonymized image and 11 other images from the second image subset was then computed using Face++'s verification algorithms [23]. One of these is the target image from the second image subset, while the remaining were randomly selected. A comparison of the similarities is shown in Fig. 1(IV). A system was built using this idea and the figure shows examples where individuals were correctly discriminated.

2.2. Privacy Enhancement with Optical Defocus

We now consider single sensors whose optical elements exhibit intentional optical defocus for privacy preservation. Unlike the k-anonymity optics discussed previously, optical defocus occurs without drawing on any on-board power source, which has advantages for miniaturization.

Optical Elements and eFOV: As in [34], we assume a distant scene which can be represented by intensity variation over the hemisphere of directions (i.e. the local light-field is a function of azimuth and elevation angles). Unlike [34], we augment the hemispherical model with a notion of scene depth, where the angular support of an object reduces as its distance to the sensor increases. We use either lensless

(I) Standard Microsoft Kinect



(II) Micorsoft Kinect with our 3D Printed Privacy Preserving Optics

Figure 3. **Privacy Preserving Depth Sensing and Motion Tracking**. We designed a 3D printed privacy sleeve that holds an off-the-shelf lens for the Microsoft Kinect V2 and that allows accurate depth sensing and motion tracking. As shown in (I), without the privacy sleeve, faces can clearly be identified in both the RGB and IR sensor images. In contrast, as shown in (II), our privacy sleeve performs optical black-out out for the RGB sensor and optical defocus for the IR sensor. Lastly, (I) and (II) also show that the native Kinect tracking software from Microsoft performs accurate depth sensing and motion tracking with and without the privacy sleeve.

or lens-based optics for defocus and, as illustrated in Fig. 5, these apply an *angular* defocus kernel over the hemispherical visual field. The range of viewing angles over which this angular support is consistent, is known as the effective FOV or *eFOV* [34]. We chose the optical elements in Fig. 5 for fabrication convenience and our theory can be used with other FOV [34, 43, 62] elements. As demonstrated by [34], every lensless element can be replaced with a corresponding lenslet element. Such an equivalent pair is illustrated in Fig. 5. In this paper, we utilize the lensless theory, even when considering lenslet systems.

The inputs to our design tool are the defocus specifications $\Sigma = \{\Delta, \sigma, R, \Theta, \rho\}$, where Δ is the angular error tolerance, σ is the desired defocus given in terms of a Gaussian blur on an image of resolution R and FOV Θ, and ρ is the length of the biggest target feature that is to be degraded by defocus blurring. For example, for a sensor designed to de-identify faces, ρ might be the size in millimeters of large facial features, such as eyes. The field of view and resolution are necessary to relate standard deviation, a dimensionless quantity, to an angular support defocus blur. The output of the tool are lensless sensor dimensions and characteristics, such as eFOV and angular support.

If we can approximate a gaussian filter of standard deviation σ by a box blur corresponding to 2σ, then, for defocus specifications Σ, the angular support is

$$\omega_o = 2\sigma \left(\frac{\Theta}{R}\right). \tag{6}$$

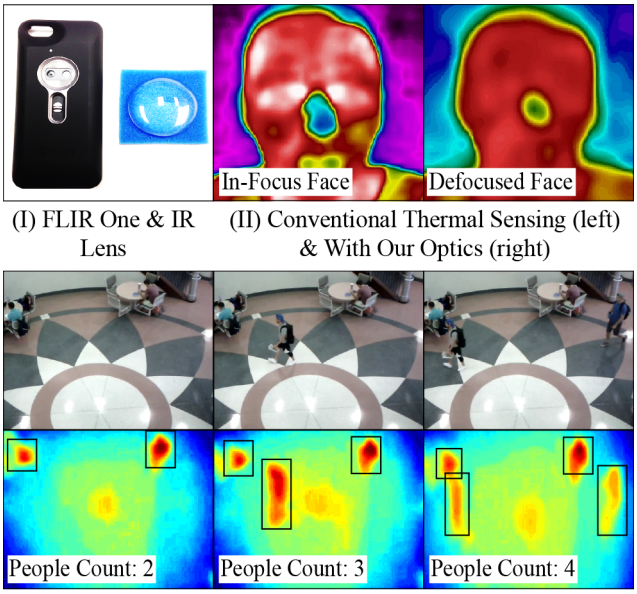
Miniaturizing a Sensor with Optical Blurring: In [34], a lensless sensor was optimally designed for maximum eFOV given an angular support ω_o and angular support tolerance Δ. We provide an additional design output, z_{min}, which is the minimum distance between the sensor and the target in order for the sensor to preserve the degree of privacy specified by the defocus specifications and it is given by,

$$z_{min} = \frac{\rho}{2tan(\frac{\omega_o}{2})}. \tag{7}$$

In summary, our algorithm takes as input defocus specifications $\Sigma = \{\sigma, \rho, \Theta, R, \Delta\}$, computes ω_o as described in Eq. 6 and applies the method of [34] plus Eq. 7 to output the optimal design with maximum eFOV, $\Pi = \{u, d, z_{min}\}$.

Example Application 1: Optical Privacy with a Time-of-flight Depth Sensor. We designed a 3D printed privacy sleeve for the Microsoft Kinect V2 that optically de-identifies faces via a defocused convex IR lens on the depth sensor and a printed cover on the RGB camera. The defocus affects the IR amplitude image while leaving the phase (or depth information) mostly intact. This occurs when the scene geometry is relatively smooth; i.e. the phasors [30] averaged by the defocus kernel are similar. The privacy sleeve as well as body tracking results under defocus are shown in Fig. 3 where the subject was 1.7m away. The angular support of the IR sensor with the sleeve was $3°$, which corresponds to lensless parameters $u = 10$mm, $d = 0.5$mm, a minimum distance, $z_{min} = 1.5$m for degrading features of 8cm and an eFOV of $64.7°$ for $\Delta = 1°$.

Example Application 2: Optical Privacy with a Thermal Sensor. We fitted a FLIR One thermal camera with an IR Lens (Fig. 4(I)) to enable privacy preserving thermal sensing via optical defocus. We performed privacy preserving people tracking by searching for high intensity blobs in the defocused thermal images Fig. 4(III). The subjects in



(I) FLIR One & IR Lens

(II) Conventional Thermal Sensing (left) & With Our Optics (right)

(III) People Tracking: Scene (top) & Thermal Image (bottom)

Figure 4. **Privacy Preserving People Tracking**. We fitted a FLIR One Thermal sensor with an IR Lens to enable privacy preserving people tracking via pre-capture optical Gaussian blurring. (I) shows the FLIR One and the IR Lens. (II) shows and image of a face taken with and without the IR Lens fitted to the FLIR One. Using this system, we were able to easily perform people tracking by searching for high intensity blobs in the optically de-identified thermal images (III).

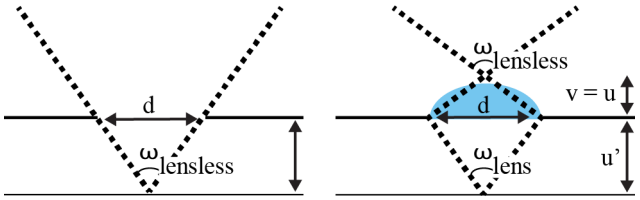
the figure were more than 5.5m from the sensor. With the fitted IR lens, the FLIR One camera had an angular support of $0.9855°$, which corresponds to a minimum distance, $z_{min} = 4.6$m for degrading features of 8cm, lensless parameters $u = 2$mm, $d = 1.29$mm, and and eFOV of $50.8°$ for $\Delta = 0.2°$.

3. Multi-Aperture Privacy Preserving Optics

In previous sections, while optical processing was used to implement privacy preserving algorithms, the actual vision computations (people counting, tracking, etc.) were performed post-capture. Here, we perform both privacy preserving and vision computations in optics by exploiting sensor arrays, which have proved useful in other domains [67].

3.1. Blob Detection with an Optical Array

A classical approach to blob detection is to convolve an image with a series of Laplacian of Gaussian (LoG) filters for scale-space analysis [40]. The LoG operators are usually approximated by differences of Gaussians (DoGs), and [34] demonstrated such computations with a single pair of lensless sensors. We build a lensless sensor array that perform both blob detection and privacy preserving defocus together. This partitions the photodetector into n sub-images with unique angular supports $\omega_{o_1} < \omega_{o_2} < ... < \omega_{o_n}$. Our prototype build with an aperture array and baffles is shown in Fig. 6. In a single shot, the sensor directly captures an image's Gaussian pyramid. When compared with a software implementation of a Gaussian pyramid, our optical array enables privacy preservation before capture. The degree of privacy afforded is directly related to the minimum angular defocus kernel ω_{o_1}. The element with the least eFOV determines the array's eFOV (although this is relaxed in the next section). Finally, the privacy preserving advantage of these arrays comes with tradeoffs; for example, the optical array provides a fixed sampling of the scale space (scale granularity) and can estimate blobs only in a fixed scale range.



(I) Lensless Configuration

(I) Lenslet Configuration

Figure 5. **Optical elements used for defocus**. We use either lensless or lenslet designs in this paper for optical defocus. The figure shows that any lenslet sensor of diameter d and image distance u can be modeled as a lensless sensor of height u and pinhole size d, and therefore we use only the lensless version in our theory.

Example Application: Privacy Preserving Head Tracking: We built a privacy preserving scale-space blob detector for head tracking. In Fig. 6 we show our prototype, which consisted of a camera (Lu-171, Lumenera Inc.) with custom 3D-printed template assembly and binary templates cut into black card paper using a 100-micron laser (VLS3.50, Versa Inc.). We divided the camera photodetector plane into nine single-aperture sensor elements using opaque baffles created from layered paper to prevent crosstalk between the sensor elements. The Lu-171 has a resolution of 1280x1024 so the photodetector array was partitioned into a 3x3 array of 320x320 pixels. Of the nine elements, three were used for our head tracking system with optical parameters $\{\Delta = 4°, \omega_{o_1} = 9.76°, \omega_{o_2} = 20.28°, \omega_{o_3} = 40.37°\}$, which corresponds to minimum distance, $z_{min} = 46.9$cm for degrading features of 8cm and an eFOV of $39.54°$. Once we detected blobs in an image, we fed the highest probability blob regions into a Viola-Jones object detector that was trained on images of head blobs moving in an office scene. The use of blobs decreased the image search area for the Viola-Jones detector by 50%. Such an example of using optics for processing reduces computation load on the system, decreasing battery usage and improving the scope for miniaturization. In the example, the head was tracked correctly in 98% of frames.

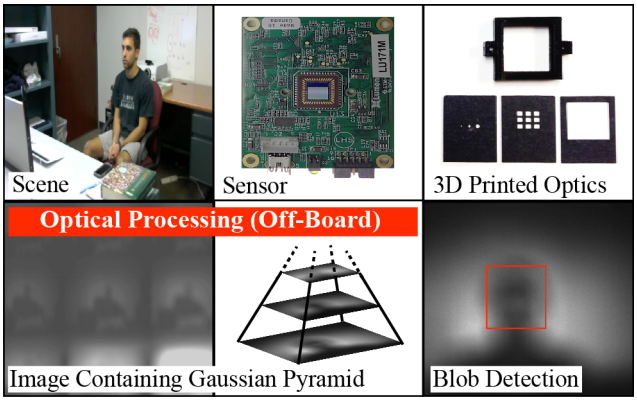


Figure 6. **Privacy Preserving Scale-Space Blob Detection.** Our privacy preserving optical blob detector uses a Lumenera Lu-171 sensor and 3D printed/laser cut optics. The sensor was divided into multiple elements, where each performs pre-capture optical defocus filtering of different aperture radii. Therefore, a single frame contains a gaussian pyramid which can be used for blob detection.

4. Miniaturizing a Multi-Aperture Sensor

In this section, we arrange optical elements within the constraints of small devices. Such packing problems have been studied in many domains [18] and the knapsack problem is a well-known instantiation [42]. We propose an optical variation on the knapsack problem that takes into account each element's angular coverage.

To see why this is needed, consider applying the traditional knapsack problem to our multi-aperture sensors. Let the total size (mass, volume or area) available for sensing optics be A. Suppose each optical element i has a field-of-view f_i and a size of a_i. Given n elements with indices $0 \leq i \leq n$, we want to find an identity vector x of length n s.t. $x_i \in (0,1)$ and $\Sigma_i x_i f_i$ is maximized whereas $\Sigma_i x_i a_i \leq A$. While this problem is NP-hard, a pseudo-polynomial algorithm $O(nA)$ has been proposed by recursively creating an $n \times A$ array M;

$$
\begin{aligned}
&M[0,a] = 0 \;\; if \;\; 0 \leq a \leq A \\
&M[i,a] = -\infty \;\; if \;\; a < 0 \\
&M[i,a] = \max(M[i-1,a], f_i + M[i-1, a-a_i]),
\end{aligned}
$$

where $M(i,a)$ contains the maximum eFOV possible with the first i elements within size constraints a and so $M(n,A)$ is the solution. Since the a_i values may be non-integers, these are usually multiplied by 10^s, where s is the desired number of significant digits. This well-known approach fails to provide the best optical element packing, because greedily increasing total eFOV does not guarantee *coverage* of the visual hemisphere. For example, a set of 5 identical elements, each having a eFOV of $\frac{\pi}{5}$, would seem to have a sum total of 180° eFOV but would redundantly cover the same angular region.

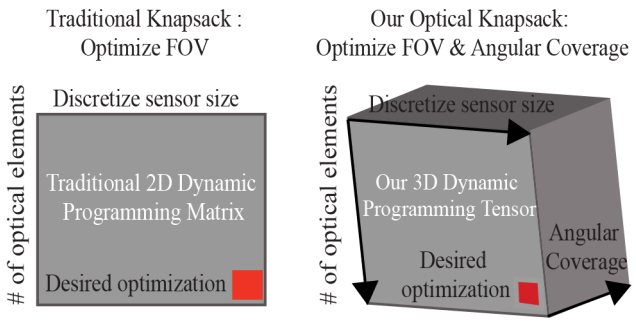


Figure 7. **Optical Knapsack Algorithm**. A traditional knapsack solution for packing optical elements might fail if the elements covered the same portion of the visual field. Our optical knapsack solution takes into account the angular coverage of each sensor and maintains the pseudo-polynomial nature of the original dynamic programming knapsack solution.

Our optical knapsack algorithm takes into account angular coverage by first discretizing the field-of-view into β angular regions, each with a solid angle of $\frac{\pi}{\beta}$. We define an array $K(n,\beta)$, where $K(i,b) = 1$ if that optical element covers the angular regions b in its field-of-view, and is zero everywhere else. We also define the array M to be three-dimensional of size $n \times A \times \beta$. As before, each entry of $M(i,a,0)$ contains the maximum field of view that can be obtained with the first i elements with a sensor of size a and $M(n,A,0)$ contains the solution to the knapsack problem. Entries $M(i,a,1)$ through $M(i,a,\beta)$ are binary, and contain a 1 if that angular region is covered by the elements corresponding to the maximum field-of-view $M(i,a,0)$ and a zero otherwise. The array M is initialized as,

$$M[i,a,b] = 0, \;\; \text{if} \;\; 0 \leq a \leq A, \;\; 0 \leq i \leq n \;\; \text{and} \;\; 0 \leq b \leq \beta$$

and is recursively updated as

$$
\begin{array}{ll}
\textbf{If } a < 0 & M[i,a,0] = -\infty \\
\textbf{For any other } a, \textbf{ for any } i & \\
\quad \textbf{If} & \\
\quad M[i-1,a,0] < & \begin{cases} M[i,a,0] = \\ f_i + M[i-1, a-a_i, 0] \end{cases} \\
\quad f_i + M[i-1, a-a_i, 0] & \\
\quad \textbf{and} & \\
\quad \sum_{1 \leq b \leq \beta} M[i-1,a,b] < & \begin{cases} M[i,a,b] = \\ M[i-1, a-a_i, b] \vee \\ K[i,b], \; b \in (1,\beta) \end{cases} \\
\quad \sum_{1 \leq b \leq \beta} M[i-1, a-a_i, b] \vee K[i,b] & \\
\quad \textbf{Otherwise } \forall b & M[i,a,b] = M[i-1,a,b]
\end{array}
$$

where $\vee$ represents the logical OR function. This optical knapsack packing algorithm adds a β multiplications and $\beta + 2$ additions to the computational cost of the algorithm. This results in a $O(nA\beta)$ algorithm, which is still pseudo-polynomial. As with the original knapsack problem, if the discretization of A and the angular regions β are reasonable, the implementation is tractable.

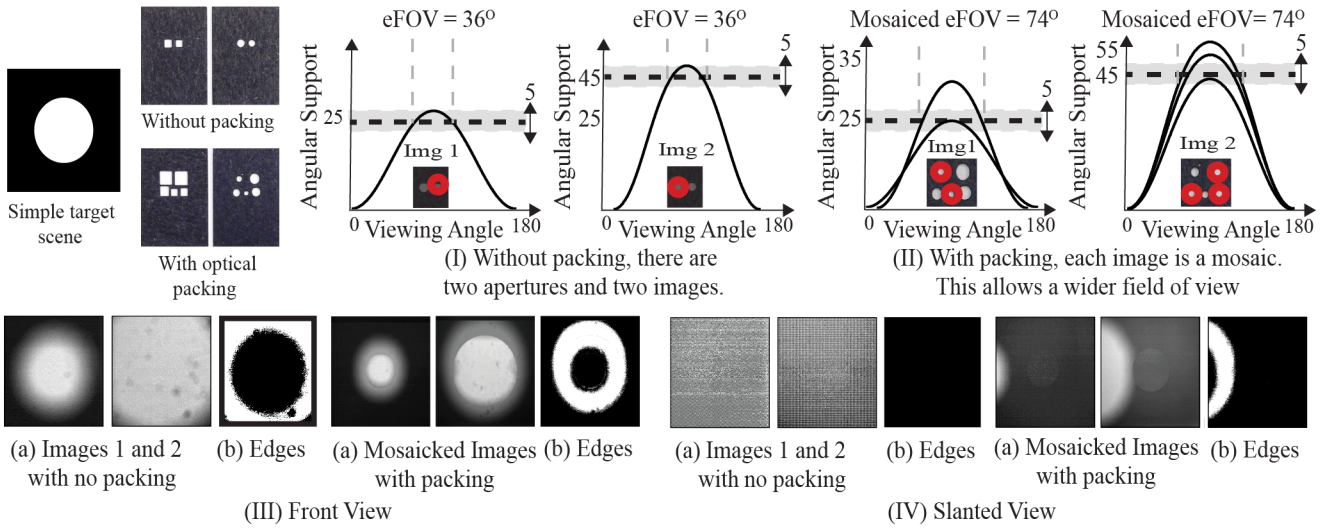


Figure 8. **Edge detection application with optical packing**. Wide angle optical edge detection has been shown [34] by subtracting sensor measurements from two different lensless apertures. [34]'s approach in (I) is *unable* to utilize the full sensor size because it requires each image to come from one sensor. In contrast, our optical knapsack technique can pack the sensor plane with multiple optical elements (II) and synthesize, in software, a wider field of view. (II) demonstrates how the angular support of multiple elements vary over the visual field, and how different measurements from multiple apertures are combined to create a mosaicked image with a larger eFOV. We perform edge detection using both the configuration from [34] and our packed sensor on a simple scene consisting of a white blob on a dark background. When the target is directly in front of the sensor (III), both optical configurations produce reasonable edge maps. At a particular slanted angle (in this case, around 15 degrees due to vignetting) [34]'s approach (IV) does not view the target (images show sensor noise) and no edges are detected. The edges are still visible for our design, demonstrating its larger field of view.

Example Application: Wide-angle Edge Detection. We demonstrate the optical packing algorithm for edge detection for a simple white disk target (Fig. 8). Our goal is two lensless sensors, each with angular supports $\omega_{o1} = 25°$ and $\omega_{o2} = 45°$ and both with error margins of $\Delta = 5°$. Fig. 8(I) shows [34]'s approach, with no packing, for a $6.6mm \times 5.5mm$ sensor and whose template height had been constrained to $u = 2mm$. Only a small portion of the sensor is used, corresponding to an eFOV of $36°$. Next we utilized our optical knapsack algorithm to maximize the eFOV on the given total area. In Fig. 8(II), a five element design is shown. Note that our algorithm only solves the knapsack part of the algorithm - the rectangular packing could be performed using widely known methods [36], but in this case was done manually. We discretized the template sizes in steps of 0.1mm and considered 30 different optical elements and discretized the angular coverage into 36 units of 5 degrees each. Since we targeted two defocus sensor designs, our 3D tensor was $30 \times 2501 \times 72$. Our dynamic programming algorithm produced the solution in Fig. 8(II), where the measurements from three elements, with aperture diameters 2.2mm, 1.9mm and 1.6mm, were mosaicked to create the image corresponding to ω_{o2} and the remaining two elements, with aperture diameters 1.2mm and 0.9mm, were used to create ω_{o1}. In the figure, the mosaicked measurements were subtracted to create a DoGs based edge detection. At a grazing angle, only the packed, wide FOV sensor can still observe the scene, demonstrating that our optimally packed design has a larger field of view.

5. Summary

We present a novel framework, which enables "pre-capture" privacy, for miniature vision sensors. Most privacy preserving systems for computer vision, process images after capture. There exists a moment of vulnerability in such systems, *after* capture, when privacy has not yet been enforced. Our privacy preserving sensors filter the incident light-field *before* image capture, while light passes through the sensor optics, so sensitive information is never measured by the sensor. Within this framework, we introduce, to our knowledge, the first ever sensor that enables pre-capture k-anonymity and multiple sensors that achieve pre-capture privacy through optical defocus. We also show theory for miniaturizing the proposed designs, including a novel "optical knapsack" solution for finding a field-of-view-optimal arrangement of optical elements. Our privacy preserving sensors enable applications such as accurate depth sensing, full-body motion tracking, multiple people tracking and low-power blob detection.

References

[1] A. M. Abdulghani and E. Rodriguez-Villegas. Compressive sensing: from compressing while sampling to compressing and securing while sampling. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 1127–1130. IEEE, 2010.

[2] P. Agrawal and P. Narayanan. Person de-identification in videos. *Circuits and Systems for Video Technology, IEEE Transactions on*, 21(3):299–310, 2011.

[3] B. Bascle, A. Blake, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *Computer Vision ECCV'96*, pages 571–582. Springer, 1996.

[4] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. K. Nayar. Face swapping: automatically replacing faces in photographs. In *ACM Transactions on Graphics (TOG)*, volume 27, page 39. ACM, 2008.

[5] M. Boyle, C. Edwards, and S. Greenberg. The effects of filtered video on awareness and privacy. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 1–10. ACM, 2000.

[6] V. Brajovic and T. Kanade. Computational sensor for visual tracking with attention. *Solid-State Circuits, IEEE Journal of*, 33(8):1199–1207, 1998.

[7] S. Browarek. *High resolution, Low cost, Privacy preserving Human motion tracking System via passive thermal sensing*. PhD thesis, Massachusetts Institute of Technology, 2010.

[8] B. H. Calhoun, D. C. Daly, N. Verma, D. F. Finchelstein, D. D. Wentzloff, A. Wang, S. Cho, and A. P. Chandrakasan. Design considerations for ultra-low energy wireless microsensor nodes. *Computers, IEEE Transactions on*, 54(6):727–740, 2005.

[9] X. Cao, Y. Wei, F. Wen, and J. Sun. Face alignment by explicit shape regression. *International Journal of Computer Vision*, 107(2):177–190, 2014.

[10] A. Chandrakasan, N. Verma, J. Kwong, D. Daly, N. Ickes, D. Finchelstein, and B. Calhoun. Micropower wireless sensors. *Power*, 30(35):40, 2006.

[11] A. Chattopadhyay and T. E. Boult. Privacycam: a privacy preserving camera using uclinux on the blackfin dsp. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[12] M. Cossalter, M. Tagliasacchi, and G. Valenzise. Privacy-enabled object tracking in video sequences using compressive sensing. In *Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*, pages 436–441. IEEE, 2009.

[13] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. The smashed filter for compressive classification and target recognition. In *Electronic Imaging 2007*, pages 64980H–64980H. International Society for Optics and Photonics, 2007.

[14] W. Dong, D. Zhang, G. Shi, and X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *Image Processing, IEEE Transactions on*, 20(7):1838–1857, 2011.

[15] B. Driessen and M. Dürmuth. Achieving anonymity against major face recognition algorithms. In *Communications and Multimedia Security*, pages 18–33. Springer, 2013.

[16] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. E. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83, 2008.

[17] F. Dufaux and T. Ebrahimi. Scrambling for privacy protection in video surveillance systems. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(8):1168–1174, 2008.

[18] H. Dyckhoff. A typology of cutting and packing problems. *European Journal of Operational Research*, 44(2):145–159, 1990.

[19] A. O. Ercan, D. B. Yang, A. El Gamal, and L. J. Guibas. Optimal placement and selection of camera network nodes for target localization. In *Distributed computing in sensor systems*, pages 389–404. Springer, 2006.

[20] A. O. Ercan, D. B. Yang, A. E. Gamal, and L. J. Guibas. On coverage issues in directional sensor networks: A survey. *Ad Hoc Networks*, 9(7):1238–1255, 2011.

[21] U. M. Erdem and S. Sclaroff. Automated camera layout to satisfy task-specific and floor plan-specific coverage requirements. *Computer Vision and Image Understanding*, 103(3):156–169, 2006.

[22] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, and T. Toft. Privacy-preserving face recognition. In *Privacy Enhancing Technologies*, pages 235–253. Springer, 2009.

[23] H. Fan, Z. Cao, Y. Jiang, Q. Yin, and C. Doudou. Learning deep face representation. *arXiv preprint arXiv:1403.2802*, 2014.

[24] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *Image processing, IEEE Transactions on*, 13(10):1327–1344, 2004.

[25] R. Fergus, A. Torralba, and W. T. Freeman. Random lens imaging. 2006.

[26] J. Gluckman and S. K. Nayar. Catadioptric stereo using planar mirrors. *International Journal of Computer Vision*, 44(1):65–79, 2001.

[27] J. W. Goodman et al. *Introduction to Fourier optics*, volume 2. McGraw-hill New York, 1968.

[28] R. Gross, E. Airoldi, B. Malin, and L. Sweeney. Integrating utility into face de-identification. In *Privacy Enhancing Technologies*, pages 227–242. Springer, 2006.

[29] R. Gross, L. Sweeney, F. De la Torre, and S. Baker. Semi-supervised learning of multi-factor models for face de-identification. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[30] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. 2014.

[31] B. Gyselinckx, C. Van Hoof, J. Ryckaert, R. Yazicioglu, P. Fiorini, and V. Leonov. Human++: autonomous wireless sensors for body area networks. In *Custom Integrated Circuits Conference, Proceedings of the IEEE 2005*, pages 13–19. IEEE, 2005.

[32] R. Horstmeyer, B. Judkewitz, I. M. Vellekoop, S. Assawaworrarit, and C. Yang. Physical key-protected one-time pad. *Scientific reports*, 3, 2013.

[33] T. A. Kevenaar, G. J. Schrijen, M. van der Veen, A. H. Akkermans, and F. Zuo. Face recognition with renewable and privacy preserving binary templates. In *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, pages 21–26. IEEE, 2005.

[34] S. J. Koppal, I. Gkioulekas, T. Young, H. Park, K. B. Crozier, G. L. Barrows, and T. Zickler. Toward wide-angle microvision sensors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12):2982–2996, 2013.

[35] S. J. Koppal, I. Gkioulekas, T. Zickler, and G. L. Barrows. Wide-angle micro sensors for vision on a tight budget. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 361–368. IEEE, 2011.

[36] R. E. Korf, M. D. Moffitt, and M. E. Pollack. Optimal rectangle packing. *Annals of Operations Research*, 179(1):261–295, 2010.

[37] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[38] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM Transactions on Graphics (TOG)*, volume 26, page 70. ACM, 2007.

[39] F. Li, Z. Li, D. Saunders, and J. Yu. A theory of coprime blurred pairs. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 217–224. IEEE, 2011.

[40] T. Lindeberg. *Scale-space theory in computer vision*. Springer Science & Business Media, 1993.

[41] G. Loukides and J. Shao. Data utility and privacy protection trade-off in k-anonymisation. In *Proceedings of the 2008 international workshop on Privacy and anonymity in information society*, pages 36–45. ACM, 2008.

[42] S. Martello and P. Toth. *Knapsack problems: algorithms and computer implementations*. John Wiley & Sons, Inc., 1990.

[43] K. Miyamoto. Fish eye lens. *JOSA*, 54(8):1060–1061, 1964.

[44] M. Mrityunjay and P. Narayanan. The de-identification camera. In *Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2011 Third National Conference on*, pages 192–195. IEEE, 2011.

[45] S. Nakashima, Y. Kitazono, L. Zhang, and S. Serikawa. Development of privacy-preserving sensor for person detection. *Procedia-Social and Behavioral Sciences*, 2(1):213–217, 2010.

[46] S. K. Nayar, V. Branzoi, and T. E. Boult. Programmable imaging: Towards a flexible camera. *International Journal of Computer Vision*, 70(1):7–22, 2006.

[47] C. Neustaedter, S. Greenberg, and M. Boyle. Blur filtration fails to preserve privacy for home-based video conferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(1):1–36, 2006.

[48] E. M. Newton, L. Sweeney, and B. Malin. Preserving privacy by de-identifying face images. *Knowledge and Data Engineering, IEEE Transactions on*, 17(2):232–243, 2005.

[49] R. Ng. Fourier slice photography. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 735–744. ACM, 2005.

[50] M. Nishiyama, A. Hadid, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi. Facial deblur inference using subspace analysis for recognition of blurred faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(4):838–845, 2011.

[51] M. Nishiyama, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi. Facial deblur inference to improve recognition of blurred faces. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1115–1122. IEEE, 2009.

[52] M. Osadchy, B. Pinkas, A. Jarrous, and B. Moskovich. Scifi a system for secure face identification. In *Security and Privacy (SP), 2010 IEEE Symposium on*, pages 239–254. IEEE, 2010.

[53] M. O'Toole, R. Raskar, and K. N. Kutulakos. Primal-dual coding to probe light transport. *ACM Trans. Graph.*, 31(4):39, 2012.

[54] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring face images with exemplars. In *Computer Vision–ECCV 2014*, pages 47–62. Springer, 2014.

[55] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1090–1104, 2000.

[56] F. Pittaluga and S. J. Koppal. Pre-capture privacy web page. *focus.ece.ufl.edu/precaptureprivacy*.

[57] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Transactions on Graphics (TOG)*, 25(3):795–804, 2006.

[58] A.-R. Sadeghi, T. Schneider, and I. Wehrenberg. Efficient privacy-preserving face recognition. In *Information, Security and Cryptology–ICISC 2009*, pages 229–244. Springer, 2010.

[59] A. P. Sample, D. J. Yeager, P. S. Powledge, A. V. Mamishev, and J. R. Smith. Design of an rfid-based battery-free programmable sensing platform. *Instrumentation and Measurement, IEEE Transactions on*, 57(11):2608–2615, 2008.

[60] S. Soro and W. B. Heinzelman. On the coverage problem in video-based wireless sensor networks. In *Broadband Networks, 2005. BroadNets 2005. 2nd International Conference on*, pages 932–939. IEEE, 2005.

[61] E. Steltz and R. S. Fearing. Dynamometer power output measurements of miniature piezoelectric actuators. *Mechatronics, IEEE/ASME Transactions on*, 14(1):1–10, 2009.

[62] R. Swaminathan, M. D. Grossberg, and S. K. Nayar. Caustics of catadioptric cameras. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 2–9. IEEE, 2001.

[63] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

[64] C. Thorpe, F. Li, Z. Li, Z. Yu, D. Saunders, and J. Yu. A coprime blur scheme for data security in video surveillance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(12):3066–3072, 2013.

[65] M. J. Wainwright, M. I. Jordan, and J. C. Duchi. Privacy aware learning. In *Advances in Neural Information Processing Systems*, pages 1430–1438, 2012.

[66] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarbotham, D. Takhar, K. Kelly, and R. Baranuik. An architecture for compressive imaging. *ICIP*, 2006.

[67] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics (TOG)*, 24(3):765–776, 2005.

[68] A. Wilhelm, B. Surgenor, and J. Pharoah. Evaluation of a micro fuel cell as applied to a mobile robot. In *Mechatronics and Automation, 2005 IEEE International Conference*, volume 1, pages 32–36. IEEE, 2005.

[69] T. Winkler and B. Rinner. Trustcam: Security and privacy-protection for an embedded smart camera based on trusted computing. In *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, pages 593–600. IEEE, 2010.

[70] W. Wolf, B. Ozer, and T. Lv. Smart cameras as embedded systems. *Computer*, 35(9):48–53, 2002.

[71] F. T. Yu and S. Jutamulia. Optical pattern recognition. *Optical Pattern Recognition, by Francis TS Yu, Suganda Jutamulia, Cambridge, UK: Cambridge University Press, 2008*, 1, 2008.

[72] H. Zhang, J. Yang, Y. Zhang, N. M. Nasrabadi, and T. S. Huang. Close the loop: Joint blind image restoration and recognition with sparse representation prior. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 770–777. IEEE, 2011.

[73] S. Zhou, J. Lafferty, and L. Wasserman. Compressed and privacy-sensitive sparse regression. *Information Theory, IEEE Transactions on*, 55(2):846–866, 2009.

[74] A. Zomet and S. K. Nayar. Lensless imaging with a controllable aperture. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 339–346. IEEE, 2006.