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, 16, 68, 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,

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

(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 $\phi$ and the sensor FOV be $\theta$. Let the minimum size of the mask that still affords the desired resolution be $M_{\text{min}}$. 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 $l_{\text{beam}}$...
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 [25] 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 \( \Delta CDE \)'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 \( \Delta CDE \)'s area. From similar triangles, we can write \( \Delta C'D'E' + C'B'BC \). From similar triangles, we write \( \Delta C' D'E \) as being created from \( \Delta CDE \) by a scale factor \( s \), and then equate the two configurations in Fig. 2 (II),

\[
\Delta CDE(1 - \frac{1}{s}) = \Delta C'B'BC. \tag{2}
\]

Consider \( \Delta CDE = \Delta COE + \Delta ODE \). From the angle-side-angle theorem, this becomes,

\[
\Delta CDE = \frac{l^2_{beam} \sin \frac{\phi}{2} \sin \phi}{2 \sin (\frac{\phi}{2} - \phi)} + \frac{l^2_{beam} \sin \frac{\phi}{2} \sin \phi}{2 \sin (\frac{\phi}{2} + \phi)}. \tag{3}
\]

Since \( \Delta AB'C' \) is a scaled version of \( \Delta ABC \), the quadrilateral area \( C'B'BC \) is

\[
\Delta ABC(1 - \frac{1}{s^2}) = \frac{M_{min} l_{mask}}{2} (1 - \frac{1}{s^2}). \tag{4}
\]

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

\[
s = \frac{M_{min} l_{mask}}{2C_1 l^2_{beam} - M_{min} l_{mask}^2}, \tag{5}
\]

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
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 \( \Delta \) is the angular error tolerance, \( \sigma \) is the desired defocus given in terms of a Gaussian blur on an image of resolution \( R \) and FOV \( \Theta \), and \( \rho \) 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, \( \rho \) 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 \( \sigma \) by a box blur corresponding to \( 2\sigma \), then, for defocus specifications \( \Sigma \), the angular support is

\[
\omega_\alpha = 2\sigma \left( \frac{\Theta}{R} \right) .
\]  

(6)

**Miniaturizing a Sensor with Optical Blurring:** In [34], a lensless sensor was optimally designed for maximum eFOV given an angular support \( \omega_\alpha \) and angular support tolerance \( \Delta \). We provide an additional design output, \( z_{\text{min}} \), 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,

\[
z_{\text{min}} = \frac{\rho}{2\tan\left(\frac{\omega_o}{2}\right)} .
\]

(7)

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

**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 \( 3^\circ \), which corresponds to lensless parameters \( u = 10mm \), \( d = 0.5mm \), a minimum distance, \( z_{\text{min}} = 1.5m \) for degrading features of 8cm and an eFOV of 64.7° for \( \Delta = 1^\circ \).

**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
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, $z_{min} = 4.6m$ for degrading features of 8cm, lensless parameters $u = 2mm$, $d = 1.29mm$, 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 proven 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 $\omega_{o_1} < \omega_{o_2} < \ldots < \omega_{o_n}$. 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 $\omega_{o_1}$. 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.

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, $z_{min} = 46.9cm$ 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 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.
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 \( \sum x_i f_i \) is maximized whereas \( \sum x_i a_i \leq A \). While this problem is NP-hard, a pseudo-polynomial algorithm \( O(nA) \) has been proposed by recursively creating an \( n \times A \) array \( M \);

\[
M[0,a] = 0 \text{ if } 0 \leq a \leq A \\
M[i,a] = -\infty \text{ if } a < 0 \\
M[i,a] = \max(M[i-1,a], f_i + M[i-1, a - a_i]),
\]

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 \( 10^a \), 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 \( \frac{\pi}{2} \), would seem to have a sum total of \( 180^\circ \) eFOV but would redundantly cover the same angular region.

Our optical knapsack algorithm takes into account angular coverage by first discretizing the field-of-view into \( \beta \) 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 region \( b \) 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,

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

and is recursively updated as

\[
\begin{align*}
\text{If } a < 0 & \quad M[i,a,0] = -\infty \\
\text{For any other } a, \quad \text{for any } i & \quad M[i,a,0] = f_i + M[i-1,a-a_i,0] \\
\text{and} & \quad M[i-1,a,b] < \\
\sum_{1 \leq b \leq \beta} M[i-1,a,b] \lor K[i,b] & \quad M[i,a,b] = \\
\sum_{1 \leq b \leq \beta} M[i-1,a-b] \lor K[i,b], b \in (1,\beta) & \quad M[i,a,b] = M[i-1,a,b]
\end{align*}
\]

where \( \lor \) represents the logical OR function. This optical knapsack packing algorithm adds a \( \beta \) 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 \( \beta \) 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 (IV) [34]’s approach 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 $\omega_{o1} = 25^\circ$ and $\omega_{o2} = 45^\circ$ and both with error margins of $\Delta = 5^\circ$. 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 $\omega_{o2}$ and the remaining two elements, with aperture diameters 1.2mm and 0.9mm, were used to create $\omega_{o1}$. 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


